Three Essays on the Empirical Market Microstructure of Money Market Derivatives

NIE, JING

Three Essays on the Empirical Market Microstructure of Money Market Derivatives

Jing Nie

Durham University Business School

Durham University

A thesis submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

March 2016

Abstract

This thesis is the first directly to study the entire limit order book of a large market. Herein, I conduct a population study on the microstructure of the Eurodollar future market, to my knowledge this is a) the first study of its type and b) the largest microstructure study ever conducted. I will build a data-drive model that incorporates information from the entire population of quotes updates and transactions on this type of future market. This thesis aims to provide a comprehensive understanding of the market microstructure on money market derivatives and the impact of high-speed algorithmic trading activity on the market characteristics and quality.

I apply a broad battery of market volatility and liquidity measurements, and gauge the proportion of high frequency algorithmic traders in the market. This thesis provides a standard asymmetric information based theoretical model to predict the relation on the term structure of Eurodollar future contracts. The prediction is a non-linear relation between the saturation of algorithmic traders (ATs) versus the impacts on the quality of the market. Therefore, I develop a novel semi-parametric estimator and model the non-linear relation between the impact of the fraction of algorithmic trading and a large set of different market quality indicators including volatility, liquidity and price informativeness. Finally, I consider the efficiency and the speed of high-frequency prices formation by implementing the return autocorrelations and vector autoregression, and also make a contribution to the trade classification algorithm using the order book data.

My findings are fourfold. First, the impact of high frequency trading (HFT) on market quality is a non-linear by implementing the semi-parametric model. This may partially explain why prior studies have found contradictory results regarding the impact of high frequency

traders (HFTs) on market characteristics. Second, prior studies only including the inside quotes or best bid best ask are limited to reflect all the information in the market. My findings suggest that the second level quoting in the limit order book is by far the most rapidly quoted element of the order book. Furthermore, I find that wavelet variance covariance of the bid and the ask side changes substantially over the term structure; providing further supporting evidence of the non-linear impact of HFTs. Finally, the adjustment time of the trade prices formation process is within one second, and the quote prices are even faster within 200 milliseconds (ms). The mid-quoted return autocorrelation is positive and gradually increase from the shortest time interval to the longest time interval. The trade prices are less sensitive to new information as the contract approaches its maturity.

Declaration

I hereby declare that this is entirely my own work unless referenced to the contrary in the text. No part of this thesis has previously been submitted else where for any other degree or qualification in this or any other university.

Acknowledgements

This thesis is for my parents who never stop giving of themselves in countless ways. Thanks for all their love and support and putting me through the best education possible. I appreciate their sacrifices and I would not have been able to get to this stage without them.

I would also like to express my gratitude and appreciation to my supervisor, Professor Julian Williams, who has shown the continuous guidance and patience, the immeasurable amount of support and encouragement, and without him I could no even begin to study in this research field.

I would like to acknowledge the help and support of the Accounting and Finance Department in the Durham University. With particular thanks to Dr Juliana Malagon and Dr Dennis Philip, whose steadfast support was greatly needed and deeply appreciated.

I would like to thank my friends Haifeng Ren, Pongsutti Phuensane, Yang Zhang, Handing Sun and Abdullah Alshami, who have been with me during the last few years.

To my great grandparents

and

my loving parents

Contents

List of Figures

List of Tables

Nomenclature

Non-Greek Conventions

a	Lowercase	Parameter
$\tilde{a}$	Lowercase	An aggregate order flow from the informed traders
$\mathbf{A}$	Lowercase	A square matrix
$\tilde{\mathcal{A}}$	Uppercase	The estimated fraction of algorithmic trading messages
B	Uppercase	The coefficient matrices
$\tilde{B}$	Uppercase	Trade intensity matrix
$\mathbf{B}$	Bold capital	The coefficient matrices of the lagged terms
$\mathcal{B}(a, b)$	Calligraphic capital	Beta distribution with shape parameters a and b
c	Lowercase	Constant matrix
$\mathcal{C}$	Calligraphic capital	Market concentration ratio
C	Uppercase	Confidence interval
C^H	Uppercase	Upper confidence bounds
C^L	Uppercase	Lower confidence bounds
d	Lowercase	Differential
d	Lowercase	Integer number
$\tilde{d}$	Lowercase	A random aggregate order flow from the uninformed trader
$\tilde{e}$	Lowercase	The aggregate order flow across both futures contracts
$\mathcal{E}$	Calligraphic capital	Pricing error ratio
$\mathbb{E}$	Hollow capital	Expectation operator
$\tilde{f}$	Lowercase	Log price function pair of future contracts
$\tilde{f}_S$	Lowercase	Log price function pair of short maturity contracts
$\tilde{f}_L$	Lowercase	Log price function pair of long maturity contracts
F	Uppercase	Companion matrix
$\mathscr{F}$	Mathscript capital	Joint density function
g	Lowercase	$1, 2$ instruments numbers
$\mathscr{G}$	Mathscript capital	Nonparametric function
h	Lowercase	Bandwidth parameter

H	Uppercase	Scaling parameter
$\tilde{H}$	Uppercase	The observed price volatility
$\boldsymbol{H}$	Uppercase	High-frequency components
$\mathscr{H}$	Mathscript capital	Haar wavelet function (mother wavelet)
i	Lowercase	A set of contract type index
I	Uppercase	A suitably sized identity matrix
j	Lowercase	Order book level, $j \in \{1, 2, \ldots, 5\}$
$\mathscr{J}$	Mathscript capital	Adjoint operator
k	Lowercase	Intraday time index
$\mathscr{K}$	Mathscript capital	Kernel densities function
l	Lowercase	The local mean of timescale length
L	Uppercase	Lag operator
L^2	Uppercase	Hilbert spaces of square integrable function
$\boldsymbol{L}$	Uppercase	Low-frequency components
$\mathcal{L}$	Calligraphic capital	Libor rate
m	Lowercase	Random-walk efficient prices
M	Uppercase	Total messages (per minute)
$\mathbf{M}$	Bold Uppercase	Marginal effect
$\tilde{\mathcal{M}}$	Calligraphic capital	Total messages (per day)
$\tilde{\mathcal{M}}^{\tilde{A}}$	Calligraphic capital	Algorithmic trading messages
$\tilde{\mathcal{M}}^{\Delta}$	Calligraphic capital	Total messages (per minute)
$\mathfrak{M}$	Mathscript capital	Mean square deviation
n	Lowercase	The number of trades
n_a	Lowercase	The number of sellers
n_b	Lowercase	The number of buyers
n_c	Lowercase	The total number of contracts on day t
n^q	Lowercase	The number of quotes
N	Uppercase	Sample size
$\tilde{N}$	Uppercase	Traders in the markets
$\tilde{N}_A$	Uppercase	Informed traders in the markets
$\tilde{N}_B$	Uppercase	Uninformed traders in the markets
$\mathscr{O}$	Mathscript capital	Operator
p, P		Trade prices
p^{-1}	Uppercase	The inverse of trade prices
p_a	Lowercase	Ask prices
p_b	Lowercase	Bid pirces
p^h	Lowercase	Highest prices
p^l	Lowercase	Lowest prices
p_m	Lowercase	Mid prices
p^v	Lowercase	Volume-weighted average prices

P^B	Uppercase	Bond prices
P^E	Uppercase	Fair price of Eurodollar future
P^F	Uppercase	Future prices
P^N	Uppercase	Notional amount
$\mathscr{P}$	Mathscript capital	Orthogonal projections operator
q	Lowercase	Trade indicator
$\mathcal{Q}$	Calligraphic capital	Quote-to-trade ratio
r	Lowercase	Return
R	Uppercase	Wavelet roughs components (residuals)
$\mathbb{R}$	Hollow capital	Set of real numbers
$\mathcal{R}$	Calligraphic capital	A range of operator
$\mathscr{R}$	Mathscript capital	Operator
$\tilde{\mathscr{R}}$	Mathscript capital	Ratio Operator
$\tilde{\mathscr{R}}_A$	Mathscript capital	Ask variance ratio
$\tilde{\mathscr{R}}_{AB}$	Mathscript capital	Bid–ask covariance
$\tilde{\mathscr{R}}_B$	Mathscript capital	Bid variance ratio
s	Lowercase	The stationary pricing error component
$\tilde{s}$	Lowercase	A tick time ($k - l < \tilde{s} \leq k$)
S	Uppercase	Wavelet smooth components
$\mathscr{S}$	Mathscript capital	Wavelet scaling function (Father wavelet)
$\tilde{\mathscr{S}}$	Mathscript capital	Spread, $\tilde{\mathscr{S}} \in \{\tilde{\mathscr{S}}^Q, \tilde{\mathscr{S}}^{Q^{1/2}}, \tilde{\mathscr{S}}^D, \tilde{\mathscr{S}}^E, \tilde{\mathscr{S}}^R, \tilde{\mathscr{S}}^{AS}\}$
$\tilde{\mathscr{S}}^{AS}$	Mathscript capital	Adverse selection
$\tilde{\mathscr{S}}^D$	Mathscript capital	Quoted depth
$\tilde{\mathscr{S}}^E$	Mathscript capital	Effective half spread
$\tilde{\mathscr{S}}^Q$	Mathscript capital	Bid-ask spread
$\tilde{\mathscr{S}}^{Q^{1/2}}$	Mathscript capital	Quoted half spread
$\tilde{\mathscr{S}}^R$	Mathscript capital	Realized spread
t	Lowercase	Daily time index
T	Uppercase	Maturity date or settlement date
T^*	Uppercase	Daily trading time length (in minutes)
$\tilde{T}$	Uppercase	Days to maturity
$\mathcal{T}$	Calligraphic capital	An exact time scaling (milliseconds, seconds, minutes, hours, days)
u	Lowercase	Error term
U	Uppercase	White noise disturbances matrix
$\mathscr{U}$	Mathscript capital	Density operator
v	Lowercase	Trade volume
$v^{1/2}$	Lowercase	Square root of trade volume
v_a, V_A		Ask volume
v_b, V_B		Bid volume

V	Uppercase	Error term matrix
$V(T)$	Uppercase	Final true value of the future
$\tilde{V}$	Uppercase	The volume of the future
$\mathcal{V}$	Uppercase	Volatility
$\mathscr{V}$	Mathscript capital	Density operator
w	Lowercase	The uncorrelated increments
W	Uppercase	Instrument variables
$\tilde{W}(.)$	Uppercase	Two-dimensional brownian motion
$\mathscr{W}$	Mathscript capital	Operator
X	Uppercase	Independent variables
Y, $\tilde{Y}$	Uppercase	Dependent variables
Z	Uppercase	Algorithmic trading proxy variables (in regression)
$\mathbb{Z}$	Hollow capital	The set of integers

Greek Conventions

α	Lowercase	Parameter; Positive regularization parameter
β	Lowercase	Parameter
$\tilde{\Gamma}$	Uppercase	The covariance matrix of the global noise
γ	Lowercase	Haar wavelet filter coefficients
$\tilde{\delta}$	Lowercase	The mark to market value
Δ	Uppercase	Difference
ε	Lowercase	Error term
$\tilde{\zeta}$	Lowercase	Global noise vectors
η	Lowercase	The orders of autocorrelation; Parameter
Θ	Uppercase	Lower triangular matrix
$\tilde{\Theta}$	Uppercase	A positive semi-definite matrix
λ	Lowercase	The number of lags
$\tilde{\Lambda}$	Uppercase	The price impact matrix (Kyle's lambda)
μ	Lowercase	Mean
ν^2	Lowercase	Wavelet variance or Incremental variance
$\tilde{\xi}$	Lowercase	Trader specific noise vectors
ρ	Lowercase	Autocorrelation function
σ	Lowercase	Standard deviation
$\sigma^{\mathbf{M}}$	Lowercase	Standard deviation of marginal effects
$\tilde{\Sigma}$	Uppercase	The variance covariance matrix of the true valuation price process
σ^2	Lowercase	Rough variance
τ	Lowercase	A time shift operator or the length of intraday time interval
$\tilde{\Phi}$	Uppercase	The variance of an individual B traders order flow

$\tilde{\Psi}$	Uppercase	The covariance matrix for each A traders signal
Ω	Uppercase	Autocorrelation
Ω^P	Uppercase	Prices autocorrelation
Ω^R	Uppercase	Return autocorrelation
Ω^S	Uppercase	Spreads autocorrelation

Abbreviations

AT	Algorithmic trading
ATs	Algorithmic traders
ATS	Automated Trading Systems
bps	Basis points
CFTC	The US Commodity Futures Trading Commission
CME	The Chicago Mercantile Exchange
COT	The Commitments of Traders
cov	Covariance operator
DAX	Deutscher Aktienindex (German stock index)
$diag(.)$	A vector of the diagonal elements of the argument matrix.
DoJ	The US Department of Justice
EBS	Electronic Broking Services
ECN	Electronic Communication Network
ED	Eurodollar
eig	Eigenvalue operator
ETFs	Exchange-traded funds
FSA	The UK Financial Services Authority
FTSE	The Financial Times Stock Exchange
FX	Foreign exchange
GDP	Gross domestic product
HFT	High-frequency trading
HFTs	High-frequency traders
i.i.d	Independent and identically distributed
IMM dates	The international money market dates
IVs	Instrumental variables
LIBOR	The London Interbank Offered Rate
LSE	The London Stock Exchange
ms	Milliseconds
NBBO	The US equity National Best Bid and Offer
NLRE	Noisy linear rational expectations
NYSE	New York Stock Exchange
OLS	Ordinary least squares

OTC	Over-the-counter
OTC-IRS	Over-the-counter interest rate swaps
PD	Positive definite matrices
RIC	Reuters instrument code
sign	Signum function
SEC	Securities and Exchange Commission
TSX	Toronto Stock Exchange
TV-VAR	Time-varying vector autoregressive model
var	Variance operator
VAR	Vector autoregression
VMA	Vector moving average
VWA	Volume-weighted average

Eurodollar Future Contracts

Eurodollar RIC [1]	Globex RIC[2]	Contracts Name with Maturity
EDH0	GEH0	March 2000/2010/2020 delivery contract
EDH1	GEH1	March 2001/2011/2021 delivery contract
EDH2	GEH2	March 2002/2012/2022 delivery contract
EDH3	GEH3	March 2003/2013/2023 delivery contract
EDH4	GEH4	March 2004/2014/2024 delivery contract
EDH5	GEH5	March 2005/2015/2025 delivery contract
EDH6	GEH6	March 1996/2006/2016 delivery contract
EDH7	GEH7	March 1997/2007/2017 delivery contract
EDH8	GEH8	March 1998/2008/2018 delivery contract
EDH9	GEH9	March 1999/2009/2019 delivery contract
EDM0	GEM0	June 2000/2010/2020 delivery contract
EDM1	GEM1	June 2001/2011/2021 delivery contract
EDM2	GEM2	June 2002/2012/2022 delivery contract
EDM3	GEM3	June 2003/2013/2023 delivery contract
EDM4	GEM4	June 2004/2014/2024 delivery contract
EDM5	GEM5	June 2005/2015/2025 delivery contract
EDM6	GEM6	June 1996/2006/2016 delivery contract
EDM7	GEM7	June 1997/2007/2017 delivery contract
EDM8	GEM8	June 1998/2008/2018 delivery contract
EDM9	GEM9	June 1999/2009/2019 delivery contract
EDU0	GEU0	September 2000/2010/2020 delivery contract
EDU1	GEU1	September 2001/2011/2021 delivery contract
EDU2	GEU2	September 2002/2012/2022 delivery contract
EDU3	GEU3	September 2003/2013/2023 delivery contract
EDU4	GEU4	September 2004/2014/2024 delivery contract
EDU5	GEU5	September 2005/2015/2025 delivery contract
EDU6	GEU6	September 1996/2006/2016 delivery contract
EDU7	GEU7	September 1997/2007/2017 delivery contract
EDU8	GEU8	September 1998/2008/2018 delivery contract
EDU9	GEU9	September 1999/2009/2019 delivery contract
EDZ0	GEZ0	December 2000/2010/2020 delivery contract
EDZ1	GEZ1	December 2001/2011/2021 delivery contract
EDZ2	GEZ2	December 2002/2012/2022 delivery contract
EDZ3	GEZ3	December 2003/2013/2023 delivery contract
EDZ4	GEZ4	December 2004/2014/2024 delivery contract
EDZ5	GEZ5	December 2005/2015/2025 delivery contract
EDZ6	GEZ6	December 1996/2006/2016 delivery contract
EDZ7	GEZ7	December 1997/2007/2017 delivery contract
EDZ8	GEZ8	December 1998/2008/2018 delivery contract
EDZ9	GEZ9	December 1999/2009/2019 delivery contract

[1]ED is the Reuters Instrument Codes (RIC) for the Eurodollar futures trading on both the open outcry platform and the Globex electronic platform.

[2]GE is for electronic trading through the Chicago Mercantile Exchanges Globex platform.

Chapter 1

Introduction

Consider a market populated by both very fast and very large (potentially slow) agents where each contract traded has a face value of one million dollars and single trades can be for 100,000 such contracts. The annual transaction volumes increase from $868 billion in 1996 to $445 trillion in 2013, and the daily trading speed drops from minutes to milliseconds over these years. In such a setting, small, but very quick, price adjustments are translated in millions of gains or losses over milliseconds investment horizons. This is a summary description of the Eurodollar Future (ED) market, possibly the most active and liquid market that remains relatively unstudied in the academic literature despite its large volume and its natural setting for algorithmic trading (AT). With trillions of dollars traded in this market over a given year, can our economy afford to leave the impact of the modern developments in trading technology on liquidity and price formation so significantly under scrutinized.

This thesis provides a comprehensive empirical analysis of the microstructure

of the Eurodollar Future market, traded on the Chicago Mercantile Exchange (CME), revolving around the effects of algorithmic trading in the quality of the market. My approach is to apply a broad battery of empirical techniques to measure the market qualities from both volatility and liquidity aspects, and combine them in a novel semi-parametric framework that studies the impact of the fraction of informed trading across a large cross section of different market quality indicators (which include liquidity, adverse selection and price informativeness). I motivate this analysis using a rational expectations model of futures trading with the differentiated capacity of traders to take advantage of foresight into the final valuation of the futures contract. Finally, I examine the speed and the accuracy of Eurodollar future prices formation at high frequency domain, and also measure the various factors influencing the price efficiency including the distance to maturity, the trading volumes, and the trading direction.

The costs and benefits of high frequency trading (HFT) activity can be succinctly summarized under the categories of execution costs/excess volatility versus liquidity provision, latency and price information process. Proponents indicate that the ability to supply liquidity by quoting and transacting at high speed improves the information processing capacity of financial markets and permit an increased volume of transactions at lower latency. Those positing a more cautious tone suggest that HFT can increase volatility, decouple bid and ask prices and harm the efficiency of price formation process, generally considered a measure of illiquidity. Recent work by Hasbrouck [2014] analyzing the quoting activity for a random sample of US equities has indicated that at the highest frequencies the correlation between the bid and ask quotes is substantially less than unity in

time frames of around 50 milliseconds (ms) and that the ratio of high frequency volatility to daily volatility can be nearly five times the level expected from a random walk with constant volatility over day. Chakrabarty et al. [2014] find little evidence of a deleterious effect of banning naked access to US equity exchanges by the SEC and cites several positives outcomes from this regulatory action. Using proprietary the US Commodities Futures Trading Commission (CFTC) data, Kirilenko et al. [2014], analyze the 2010 Flash Crash that appears to have started in E-mini S&P 500 futures, illustrate that whilst HFT was not the trigger of this crash, high frequency traders (HFTs) do appear to have exacerbated the degree of price volatility. A number of recent theoretical contributions have come out against HFT, indicating that the benefits of high speed transactions are outweighed by the negative effect of a subset of traders having substantially greater opportunities simply through the speed of access, see Kyle and Obizhaeva [2013] and Foucault et al. [2015].

In contrast, Hendershott et al. [2011] illustrate empirically that the introduction of the 'autoquote' facility for US equities has a significant and positive effect on a number of liquidity proxies (positive in the sense of an increase in liquidity). Further support for HFT as a net provider of liquidity maybe found in Hendershott and Moulton [2011] and Riordan and Storkenmaier [2012]. In each case, the statistical experiment is founded on the introduction of a new technology providing new avenues for HFT. Castura et al. [2010] utilize the variance ratio method on the Russell 1000 and 2000 markets and illustrate that the high-speed trading has benefits to the market efficiency. Chaboud et al. [2014] find that the high-speed algorithmic trading can limit arbitrage opportunities and improve the

price formation process in the foreign exchange (FX) markets. Grünbichler et al., 1994, Hendershott and Riordan [2011] and Chaboud et al. [2014] provide evidence to support that HFT can enhance the price formation process. For various markets and asset classes, Harris [1989], Stoll and Whaley [1990], Chan [1992] and Huang and Stoll [1994], Engle et al. [2012], Frino and McKenzie [2002] and Menkveld and Zoican [2014] each either finds evidence to suggest that HFT improves market characteristics or finds no evidence that the level of market quality is compromised by the presence of HFT or its introduction.

The common microstructure studies focus on the equity market and foreign exchange market due to the data availability. Give the lack of coverage in the literature, this thesis conducts a population study on the microstructure of the three-month Eurodollar future market. Eurodollar future contracts is a London Interbank Offered Rate (LIBOR) referenced contract with a 10-year life span. The most popular Eurodollar contracts are quarterly traded, via both open outcry and Globex electronic platforms, such as the March, June, September or December delivery contracts. I have constructed two unique quarterly Eurodollar future datasets. One is the inside quotes and transaction data with every update from 1996 to 2013; the other is the entire messages within the limit order book from 2008 to 2014.

The reason I choose this interest rate derivatives market, first of all, it is the largest future market in the world with the high level of liquidity and fast updating speed [Commodity Futures Trading Commission, 2012]. Each Eurodollar futures contract has a notional value of $1 million. Especially, after the Dodd-Frank Act released in 2010, it sets the margin requirements on the interest rate swaps to

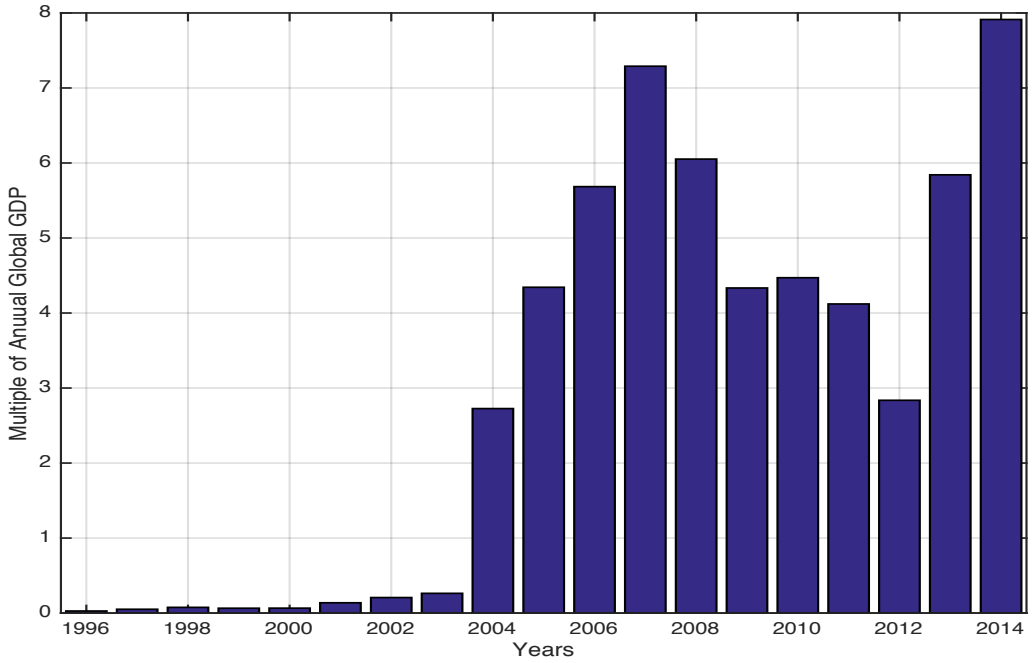


Figure 1.1: Annual Eurodollar Future Trades Volumes versus Global GDP

Notes: This figure represents the ratio of the annual Eurodollar futures trading volumes and the annual global GDP from 1996 to 2014. The 40 Eurodollar future contracts data is obtained from CME Exchange tapes via Thomson Reuters Tick History database. For each calendar year, the ED trades volumes are calculated as the $\sum V_i$, where V is the trades volume and i is a contract type index. The 1996–2014 annual global GDP data is collected from the World Bank World Development Indicators via Datastream.

lower risks in the trading system. This action impels more traders to move to Eurodollar futures because it is now commonly cheaper to construct the floating leg of interest rate swaps using Eurodollar strips. Figure 1.1 depicts the annual Eurodollar future trading volume and the annual global gross domestic product (GDP) from 1996 to 2014. We can see, before 2004, the annual volumes of Eurodollar futures are less than the annual global GDP. Once it passes 2004, the annual ED trading volume increases from $118.62 trillion in 2004 to $419.41 trillion in 2007. The ratio of ED volumes and the global GDP shows a dip

between 2008 and 2012, which is in keeping with the period of recent financial market stress. The 2013 global GDP recovers to $445.39 trillion, and then the annual ED trading volume in 2014 reaches about eight times as many as the global GDP[1]. Another reason I choose this future market is that it has a unique international money market (IMM, legacy terminology) dates, so I can study the term structure of Eurodollar futures based on the length to maturity instead of calendar dates.

1.1 Main Empirical Research Questions

The impact of high frequency trading activity on the characteristics of financial markets is a topic of considerable contemporary interest. To fill the identified research gap on the microstructure of Eurodollar future market, the main research questions under investigation in this thesis are

- *In a continuous electronic central limit order book, how does the presence of high frequency trading impact market quality?*
- *In the absence of data on the specific types of accounts, or the ability to track quoting activity from a specific account, is it possible to construct a meaningful measure of high frequency trading in a continuous electronic central limit order book?*

[1]The ratio between ED volume and global GDP in 2014 only includes the first seven months, because my Eurodollar future transaction data is gathered from January 1996 to July 2014. The annual global GDP is from the World Bank World Development Indicators via Datastream database.

- *Can we design a model of futures trading that matches the observed pattern of market liquidity and market quality?*
- *Can we use these measures to influence the design of markets and better inform market participants of all types on how best to trade in such a market?*

To answer these questions I have developed a research programme that uses theoretical models of futures trading which is then used to motivate a series of empirical models that provide evidence on the association between market quality and high frequency trading.

By market quality I refer to two aspects: the level market volatility and the degree of market liquidity provision. I will carefully identify these two components in both Chapter 2 and Chapter 3.

Chapter 2 investigates how do the Eurodollar future prices behave and determine the impact of high speed quoting on very short-term volatility. In this chapter, I analyze the microstructure of the volatility of inside quotes in the Eurodollar futures market by implementing the wavelet techniques to extract the local averaging effects to calculate the variance and covariance of the order flow over days at various timescales. The variance ratio records the changes of quadratic price variation at different time scales, usually relative to the daily variation. The covariance ratios provide some insights into the synchronization and the agreement between the bid-side and ask-side. If the Eurodollar future market is an efficient market and both bids and asks prices follows a Brownian motion, the theoretical prediction of the variance ratio should follow a linear relation and close to one.

This chapter finds the high variance scaling and low covariance ratio at short timescales, which suggests the existence of high-frequency microstructure noises and the disagreement between ask and bid side at highest frequency in this market. Besides, as a Eurodollar future contract approaching its maturity, the quote prices become more volatile.

Chapter 3 theoretically analyzes the relation between informed traders and the quality of futures markets. I demonstrate theoretically that the anticipated marginal effect of additional HFT activity is non-linear. Using the entire Eurodollar future limit order book and trades data, this chapter also measures the proportion of high frequency algorithmic trading as the HFT proxy and calculate a battery of market liquidity indicators, including bid-ask spreads, quoted depth and adverse selections. The results illustrate the changes in both the market liquidity and high-speed trading behavior over the term structure of the Eurodollar contract. I also detect that the quotes at level two from the order book provide the most high-speed active with a large volume of quotes.

Chapter 4 empirically captures the relationship between HFT and indicators of market quality using the Eurodollar futures market as a case study. I estimate a semi-parametric model with HFT intensity, as the proportion of total market activity, used as a predictor of 'market quality' which is a loose description for a series of indicators of liquidity, execution risk and adverse selection. I demonstrate, empirically, that a critical saturation point, where the contribution of an extra percentage point of HFT no longer reduces market quality but actually increases, is a) supported by empirical observation and b) relatively stable across a variety of indicators.

Chapter 5 measures the efficiency of price formation process and the trading mechanism and answers how fast you need to be in the Eurodollar future market. In this chapter, I undertake a comprehensive analysis of the price discovery and price efficiency at a very high-frequency domain. To assess the ability and the speed of the prices incorporating information, I use the return autocorrelations at 26 different time intervals, from tick-by-tick level to 30 minutes, to measure the price efficiency. Then I capture the pricing error, the deviation of transaction prices from the efficient price, using the vector autoregression model. What I find is the return autocorrelations are positive and gradually increase from the shortest time interval to the longest time interval. The findings also suggest the adjustment time of trade prices is in a very short time frame within one second. The quote prices adjust even faster than the trade prices, in the milliseconds frame. Moreover, as the Eurodollar future approaches its maturity, the trade prices are less sensitive to incorporate any information on the market.

1.2 Contribution of This Thesis

This thesis makes some distinct contribution to the knowledge and understanding in the market microstructure research area. First of all, this research is the first one to conduct a population study on Eurodollar future market, and to my knowledge, it is the largest microstructure studies at this time. Every message in the Eurodollar future market has been collected. My historical Eurodollar future data spans 19 years, from 1996 to 2014, including tick-level quotes and trades at milliseconds timestamp. So I am one of the first researchers to directly model and

analyze the limit order book for this type of market. The second contribution of this thesis is that it captures a wide variety of price oscillations utilizing a recently developed wavelet techniques to scale the prices volatility. Although it has been adopted by Hasbrouck [2014] in the equity market, this signal processing technique is the first time used to analysis the interest rate derivatives.

From the point of view of exchange providers, CME Group [2010] makes use of the requirement for registration of algorithmic traders (ATs) connections to the Globex trading platform and runs a simple linear analysis of the impact of the level of HFT activity on bid-ask spreads finding evidence to support the conjecture that greater HFT leads to deeper markets. So how to reconcile these opposing empirical and theoretical findings? A first pass is to look at the theoretical structure of the market, when a set of traders has the ability to effectively predict future outcomes by speed alone. I will show that in a relatively standard, but realistic setting that as the number of traders with this trading advantage increases, the size of the price impact for a given level of order-flow imbalance, whilst initially increasing eventually decreases and then tends monotonically to zero. The ability of a market to find the efficient price in the presence of a given degree of order-flow imbalance is the definition of market depth. Hence, the theoretical prediction will be that for a given number of ATs generating HFT order flow, the negative effect on liquidity will only be significant when a small number are in the market exploiting their ability to effectively 'front run' the efficient price through their speed advantage. The idea that as the number of ATs increases their effect cancels is more-or-less obvious, however, my contribution is specifically on the shape of the adjustment and I will demonstrate that this

'shape' is supported by observation from empirical evidence.

Another contribution focuses on my novel empirical approach to analyzing the impact of HFT and the number of ATs within my chosen market. Motivated by the theoretical findings, I determine the fraction of ATs within a market by looking directly at the speed of update of the complete limit order-book. I utilize the Eurodollar futures market for this exercise as it is a large liquid market with a very high level of activity. The information flows to this market from the LIBOR reference rate are relatively well understood and I have access the full and uncensored limit order book. I classify quotes with update speeds of less than 25 milliseconds as being HFT generated by an AT, which I find on average to be around 65% of the total number of messages, on average. The fraction of messaging volume generated at this speed is then used as a proxy for the intensity of HFT activity as a fraction of total messaging. I will show that 25ms generates a distribution of HFT within the market that is in line with the exchanges own assessment of the level of activity given their knowledge of the actual account numbers and hence whether order flow originated from the active account connection of an AT or from a terminal driven by registered individual.

I then regress this intensity measure HFT from ATs as a fraction of messaging semi-parametrically onto a variety of measures of liquidity and adverse selection utilizing a variety of instrumental variables to assist in identification. I show that the pattern of the effect of the increasing fraction of HFT from ATs on these liquidity metrics follows the pattern established in the theoretical model very well (albeit that my theory model predicts a slightly lower saturation point than empirically indicated by several metrics). I believe that my results go some way

to answering the question of why many prior studies using linear models have found substantially differing results.

Furthermore, I develop a new effective algorithm to classify the trades directions. Because the traditional Lee and Ready [1991] algorithm does not work when the timescale is smaller than 1 second, I develop a unique volume weighted average trade classification algorithm to replace the Lee-Ready method, specially designed for the microstructure studies using the limit order book data at high frequencies. This new technique has been employed to calculate a range of market liquidity spreads, spread autocorrelations and short-term pricing errors in Chapter 5. After comparing the new algorithm with the Lee-Ready algorithm using a toy example, the traditional method can only classify about 33% of trades as a buy-side trade or a sell-side trade, but the volume-weighted algorithm can classify about 96% among all the trades. The other 4% mainly belongs to the iceberg trades, which are only reported with prices but hidden their volumes.

Two final important take-home observations specific to futures markets are a) that studies that ignore high levels of the order-book than level one miss a great deal of activity, level two prices across most maturities are the most actively updated by a substantial margin. My empirical observations indicate that level two quoting is a common domain for ATs and the recipient of a considerable fraction of HFT activity. b) the maturity of the contract is non-linearly correlated with the numbers and speed of updating. Long maturity contracts have up to six orders of magnitude less quotes than those with around three to four years of maturity and there is an asymmetric drop in activity across the order-book as the tenor of the contract enters its final year. Taken together, any microstructure

analysis of a futures market will need to sample quite widely across maturities to gain a full picture of how order-flow and more pertinently speed of order-flow impacts liquidity.

1.2.1 Summary of Contribution

1. Formulation, solution and simulation of a new multivariate rational expectations model specifically formulated for futures contracts. This is implemented in Mathematica.

2. Formulation of a new semi-parametric empirical model of market quality, including the introduction of a new bootstrapping approach to confidence bounds for Gaussian partially linear semi-parametric models with semi-parametric instruments, implemented in Matlab.

3. The design of a new technique for imputing the trade direction from an electronic limit order book.

4. Specification, analysis and simulation of a new linear vector autoregression model of market quality.

5. The design of new algorithms for each of the above empirical models to deal with multi-terabyte data.

6. The design of new infrastructure to process and match large scale databases from the raw central limit order book of catalogued exchanged tapes.

7. First research to analyze the complete limit order book.

8. First thesis to conduct population-level studies on a particular market (all quotes and trades across the order book from 2008 to 2014).

The remainder of this thesis is organized as follows. Chapter 2 analyzes the microstructure of the volatility of quotes using the Eurodollar futures market as a case study. Chapter 3 proposes a non-linear theoretical prediction in terms of the effects of informed traders, and then measures the a set of market liquidity indicators and the fraction of high-speed algorithmic trading in this market. Chapter 4 employs a semi-parametric approach and captures the non-liner marginal effects of HFT on the quality of Eurodollar future market. Chapter 5 examines the efficiency of price formation at a very high-frequency domain. I also provide an extensive online appendix at https://www.dropbox.com/s/uxbilz8879azdkb/Thesis_OnlineAppendix.pdf?dl=0, which contains a comprehensive series of complementary information to support this thesis.

1.3 Contents of Online Appendix

This online appendix provides complementary analyses to the main tables and figures in this thesis. It contains the following content to support the analysis contained within this thesis:

Appendix A contains additional details and supplemental results for Chapter 2 *High speed quoting and excess volatility: evidence from money market derivatives*, including the summary statistics tables, additional quotes volatility tables and wavelet quotes variance plots.

- Appendix A.1 shows the descriptive statistics of quotes (both inside quotes and entire limit order book) and transactions for all Eurodollar contracts (see Table A.1 – Table A.4).

- Appendix A.2 provides the median results of bid variance, ask variance, covariance and correlations at all timescales (see Table A.5 – Table A.20).

- Appendix A.3 includes other types of wavelet variance-covariance results, including bootstrap mean, mean, max, min and standard deviation (see Table A.21 – Table A.60).

- Appendix A.4 presents year-long median value of the daily bid-ask wavelets variance, covariance and correlations (see Table A.61 – Table A.98).

- Appendix A.5 shows the estimation results of high speeding quotes. Table A.99 –Table A.102 contains the average high speed quoting results from top 10% days with the highest number of bids. Table A.103 – Table A.106 provide the supplementary tables for variance ratio and timescales of the most active day.

- Appendix A.6 depicts the wavelet decomposition with bids and asks variance ratio for all 9 timescale levels. Here I present wavelet variance plots for the March, June, September and December delivery contracts groups (see Figure A.1 – Figure A.4) and wavelet variance plots for each Eurodollar contract (see Figure A.5 – Figure A.44).

Appendix B is complementary to the Chapter 4 *High frequency trading and execution risk dynamics: evidence from eurodollar limited order book data.* This

section presents a significant number of ancillary treatments for the robustness check, including marginal effects plots with algorithmic trading proxies at various thresholds and different instrumental variables.

- Appendix B.1 shows some examples of the impacts of AT proxy thresholded at 200 ms on the execution risks by the implementation of the conventional Robinson estimator using the *semipar* command in Stata 13. These plots from Stata provide complementary information for the robustness of the marginal effects plots from Matlab in the thesis (see Figure B.45 – Figure B.61).

- Appendix B.2 presents the marginal effects of algorithmic trading proxies on the execution risk dynamics utilizing the partially-linear semi-parametric model. Figure B.62 to Figure B.71 show the marginal effects of AT proxy thresholded at 25 ms. Figure B.72 to Figure B.80 show the marginal effects of AT proxy thresholded at 200 ms.

- Appendix B.3 illustrates the robustness check for the marginal effects of algorithmic trading on market quality indicators. For the robust check, I employ two different instruments instead of the total messages. Firstly, I use the number of traders as the instrument for the proxy of algorithmic traders thresholded at 25ms (see Table B.107 and Figure B.81 to Figure B.97). I also use the number of traders as the instrument for the proxy of algorithmic traders thresholded at 200ms (see Table B.107 and Figure B.98 to Figure B.114). Alternatively, I also use the volume to message ratio as the instrument for the proxy of algorithmic traders thresholded at 25ms

(see Table B.109 and Figure B.115 to Figure B.131); and the results and plots of the marginal effects of AT proxy thresholded at 200ms using the volume to message ratio as instrument are in Table B.110 and Figure B.132 to Figure B.148.

Appendix C provides additional supplemental results for Chapter 5 *High-frequency price discovery and price efficiency on interest rate futures*. This section contains the return autocorrelation, liquidity spreads, spreads autocorrelation and regression results tables to support the analysis of Chapter 5.

- Appendix C.1 illustrates the mid-quote return autocorrelations from 30 days before maturity subsample to 1800 days before maturity subsample (see Table C.111 – Table C.120).
- Appendix C.2 depicts the daily first-order, fifth-order and tenth-order autocorrelations of the mid-quote returns at various time intervals sorted by the days to maturity across 40 Eurodollar contracts (see Figure C.149 – Figure C.152).
- Appendix C.3 describes the average and median of a range of market liquidity spreads of the Eurodollar future market, including bid-ask spreads ($\tilde{\mathscr{S}}^{Q}$), quoted half-spread ($\tilde{\mathscr{S}}^{Q^{1/2}}$), quoted depth ($\tilde{\mathscr{S}}^{D}$), effective half-spread ($\tilde{\mathscr{S}}^{E}$), realized spreads ($\tilde{\mathscr{S}}^{R}$), and adverse selection ($\tilde{\mathscr{S}}^{AS}$). This section includes all the spreads from 30 days before maturity subsample to 1800 days before maturity subsample (see Table C.121 – Table C.130).
- Appendix C.4 contains the first-order, fifth-order and tenth-order autocor-

relations of liquidity spreads with the 30 days before maturity subsample to the 1800 days before maturity subsample (see Table C.131 – Table C.160).

- Appendix C.5 represents the daily first-order, fifth-order and tenth-order autocorrelations of a set of liquidity spreads sorted by the days to maturity across 40 Eurodollar contracts, including quoted half spreads, realized spreads and adverse selection (see Figure C.153 – Figure C.161).

- Appendix C.6 shows the impulse response plots of five variables, including the trade direction, trade volume, square root of trade volume, and asks and bids market concentration for the 10-year full sample period, one-year tenor group, three-year tenor group, and five-year tenor group (see Figure C.162 – Figure C.171).

- Appendix C.7 summarizes the multivariate linear regression results of a set of market activities indicators, price discovery factors, mid-quote return autocorrelations, market liquidity spreads and spreads autocorrelations, within the 10-year full sample period, one-year tenor group, three-year tenor group, and five-year tenor group (see Table C.161 – Table C.184).

Chapter 2

High Speed Quoting and Excess Volatility: Evidence from Money Market Derivatives

The futures contract on the 3-month Eurodollar Deposit is one of the most actively traded in global financial markets. Most reliable estimates suggest that the notional outstanding on these contracts makes this the worlds largest futures market.[1] *Several regulatory interventions into the probity of submissions of the Eurodollar futures reference rate, the 3-month LIBOR, have focused attention on this market, and in particular its transactional microstructure and information content. In this study, I analyze the microstructure of the volatility of quotes in the Eurodollar futures market by implementing the recently developed approach of Hasbrouck [2014]. Using a comprehensive database of every update to the inside-quotes – some two billion observations, from 1996 to 2014 – my results paint a fascinating picture of the changing behavior of traders in these very actively traded contracts. For long maturities the variance ratios of the updates of the inside-quotes are well behaved and approach unity – this is in contrast to the equity market. However, as the contracts approach maturity the number of updates to the inside-quotes increases dramatically (by six orders of magnitude for the one year prior to maturity) and the variance ratios now begin to approach those found in the equity market literature; finally, as the contract approaches maturity, this effect then declines dramatically. My results lend credence to the theory that maturity effects are driven by speculative HFT traders entering the futures markets for only certain epochs of the futures lifecycle.*

[1] Source: CFTC [Commodity Futures Trading Commission, 2012, Barclays Settlement].

Empirical research on the microstructure of interest rate derivatives is a relatively overlooked area of financial economics. However, recent civil and regulatory actions against several large financial institutions with substantial activities in interest rate related swaps and futures has shone a light on these very actively traded, but little studied products. In this chapter, I conduct an extensive analysis of the 3-month Eurodollar futures market and in particular the short-term, intraday, volatility of quotes over the lifecycle of the futures contract. I will utilize techniques to extract the 'variance and covariance ratios' of noise components at various timescales over a single day and illustrate the marked change in the behavior of the market near the maturities of these contracts. Variance ratio analysis documents the change in the expected level of quadratic variation at different time scales and maybe used to determine the optimal sampling frequency for variance-covariance analysis and determine the impact of high speed quoting on very short-term volatility. Additionally, the analysis of the covariance ratio between the inside-bid and inside-ask, provides insight into the time-scale at which there is a breakdown in the agreement between the buy and sell side of the order book, an important measurement of market liquidity and the informativeness of quotes.

My evidence points to a marked maturity effect decoupling the inside-bid and insides-ask prices, resulting in sudden changes in liquidity even when the contracts are actively traded. An implication of the results is that these frictions may allow sophisticated traders with quick reactions to exploit informational advantages near maturity. The results indicate that whilst abnormal variance ratios appear near maturity, the clustering effect usually occurs prior to the last day of trading and is coincidental with high levels of market activity. The market for the 3-month Eurodollar futures contract (henceforth Eurodollar Future), when measured by the notional value of open interest is the "most liquid and largest" [Commodity

Futures Trading Commission, 2012, p. 5]. The CFTC reports that the notional open interest was in excess of $564 trillion in 2011. My dataset is comprehensive in that I am able to analyze every quote at the inside-spread from 1996 to 2014. The volume of this market is quite extraordinary, with actual traded volumes across 40 contracts averaging in the trillions of dollars over my sample period – indeed the dollar volumes of quotes are in the quintillions of dollars. At any given time the Eurodollar market has notional amounts outstanding that are around an order of magnitude more than global GDP. With each contract having a notional value of one million dollars, small relative price fluctuations can represent very large absolute dollar amounts.

My approach is to assess the relative quadratic variation for every millisecond timestamped updates of the best bid and offers (the inside-spread), in this market and compute variance ratios at different timescales for the bid and ask. The time variation in the variance ratios provides a guide to the mechanisms driving the update of prices over a given day. In an important contribution, Hasbrouck [2014] has outlined an empirical approach to identify quote-induced volatility in equity markets. The importance of this topic cannot be understated as "[bid-offer] volatility degrades the informational content of the quotes, exacerbates execution price risk for marketable orders, and impairs the reliability of the quotes as reference marks for the pricing of dark trades" [Hasbrouck, 2014, p. 1:2]. Hasbrouck [2014] utilizes equity best bid and best ask data from 2001 to 2011 and associates this with the transition to electronic trading. In keeping with the equity market, futures are traded through a limit order book, however, futures contracts are centrally cleared by the futures clearing house, so there are subtle differences that may play a role in differentiating the results found herein and for equivalent data from the equity market.

Given that the pattern of high frequency trading is closely correlated to the maturity of the futures contract, I impute the term structure of quoting activity by type in the market and demonstrate that activity varies in relation to the IMM dates. Given that the Eurodollar future has a very important relationship to the interest rate swaps (IRS) market the dynamics of execution costs have significant implications for the overall structure of the dollar fixed income market. Indeed, it is instructive at this stage to provide a short example of the type of problem I address. Let us consider a two-year interest rate swap with the present value of floating payments and fixed payments equalized to be $49.967 billion. For convenience, I will assume that the floating payments are referenced to the same IMM reset dates which are the quarterly maturities for the Eurodollars, hence the basis risk is zero.[1] Consider a futures strip that hedges the floating rate payments, the holder will periodically need to buy (to hedge fixed rate payers exposed to the risk of falling rates) and sell (to hedge fixed rate receivers exposed to the risk of rising rates). Let us consider a standard quarterly rebalancing, several tools exist that can be used to compute the convexity adjustments.[2] I will focus just on the economic implications of transactions on the floating side, when rates are

[1] I should note that this is a relatively small position compared to some implied by the trades in my data-set, I use this toy example as the number of contracts needed to compute the hedge is relatively trivial to calculate, see 'CME Group – Interest Rates and Eurodollars' for other extended examples. Why is Eurodollar notional position so large and getting larger? Following the Dodd-Frank Act of 2010, Part 39, Subpart B, Section 39.13(2) (ii) which covers margin requirements indicates that futures contracts must be covered by performance bonds with one-day liquidation and interest rate swaps require a five-day liquidation timetable. The Dodd-Frank Act has, as such, substantially increased margin costs for the previously relatively unregulated bilateral swaps market. The result is that a ten-year Eurodollar futures contract constituting a portfolio of futures will have a significantly lower margin requirement than an equivalent interest rate swap, as the futures are centrally cleared and marked to market daily. As of September 2014, a proposal for bilateral un-cleared swaps to be cleared by a 'qualified-central-clearing-party', the CME 'ClearPort' system is an example of this. Therefore, the swaps will in effect mimic many of the properties of futures.

[2] Bloomberg has the EDS functionality, CME provides the 'Swap Equivalents' tool and in Matlab the `LiborMarketModel` function can be used to accomplish the same goal to nest in more automated platforms.

falling, such as in the post-2008 period although similar counter examples can be constructed on the buy side.

The floating rate counterparts will sequentially sell Eurodollars at each of seven rests over the life of the swap. The largest trading volume in my sample period from 1996 to 2014 from the viewpoint of completed trades in a specific contract is EDH1 on February 18, 2011 (contracts maturing March, 2011), with $49.967 billion of ED futures contracts were traded. It should be noted that the lack of convexity between price and yield exhibited by the Eurodollar curve means that the hedging ratio will change slightly per time, so the textbook approach does not quite hold, a fact that has caught many trainee swap traders. Consider a single quarterly sale of 49,967 Eurodollars, the notional size of $1 million, this will be one quarter prior to maturity. If the inside-buy is currently trading at 99.66 and the inside-sell is 99.6775, with more than 50,000 contracts available at both the sides of the spread (these are real spreads and trades from February 18, 2011 prices for March maturing contracts) and the trader decides to cross the spread, he expects to complete the sale at or above 99.6775. This is under the presumption that in the 35 to 70 milliseconds required to execute the best buy does not move substantially. However, my variance ratios indicate that shortly before maturity dates the volatility of inside-bid substantially higher than that expected under a daily level of variance. At execution for a large trade such as this one, the trader takes a risk that either the best bid will suddenly withdraw and the trade is not completed. Note that CME has four main types of order, none of which is a market order in the traditional sense.[1] Or alternatively the

[1]These are: the standard limit order that posts itself to a level in the order book concomitant with the price within the quote and if it is a bid or an ask (or creates a new level if there are no matching existing prices) relative to the best-bid or best-ask; a market order with protection, effectively an order that buys(sells) any contracts the best ask(bid) price plus a specific spread and if there is any remainder posts the price and remaining volume as a new limit-order at the limit of the range; a market to limit, which is effectively the same as the market order with

best bid jumps and a high frequency trader buys the contracts at 99.66 and instantly sells at 99.6775 (the previous best ask) realizing 1.75 basis points per contract, or instantly gaining $2,186,056.25, this is a relatively benign case, as I have measured the intraday kurtosis of the inside-spread quote by quote to be high as 830,000.

This might simply be considered unlucky in that the trader happened to choose a trade as the market moved, however the covariance/correlations scaling ratios tells us a different story. Correlations between the inside-bid and inside-ask decay dramatically at the quote level, indicating that the price ceiling on the ask side is highly uninformative of the overall order book. I find that the effects are related to heightened trading activity and this is high correlated to specific maturity dates in the contract cycles. This appears not to be a coincidence as the IMM dates coincide with a great number of swaps seeking to reduce basis risk, given that prevailing interest rate conditions are well understood net buying and selling pressure can be anticipated relatively easily. Whilst in the equity market, variance ratios and bid-ask decoupling can amount to potential HFT arbitrage profits of a few hundreds of dollars a trade, the Eurodollar market permits a far quicker root (in terms of transactions) to large HFT profits.

The contribution of this chapter is twofold: first, to extend the approach to capture a wide variety of price oscillations; and second, I am the first to conduct this analysis on interest rate futures priced relative to the LIBOR, something

protection except the upper range is the best-bid best-ask; a resting stop limit is an order that sits until a trigger price is traded and then a limit order is submitted to the market, again if the order cannot be filled immediately then the price within the stop limit is posted; finally, a stop limit with protection is the same as a stop limit, but the triggered order has a protected range. For each order type there are a variety of duration and display qualifiers, such as the standard fill-and-kill (if part filled the remainder is killed), good 'till cancelled (order needs to be cancelled by the trader) and more intriguingly the MaxShow or IceBerg, which masks part of the volume of the quote, as each chunk of the quote is filled a new identical quote is displayed.

that is unique in the literature. My main objective is to disentangle systematic maturity effects from 'other' sources of microstructure noise, such as possible informed trading. I therefore construct an analysis that follows Hasbrouck [2014], however I decompose the average variance-covariance ratios relative to maturity date and activity to illustrate the systematic effects of high speed quoting on short time-scale volatility. My results can be roughly summarized by the timeline in Figure 2.1, with three main 'epochs' for the contract – these of course also relate to the contracts position on the implied corresponding forward rate curve.

The first epoch, begins at inception the equivalent of a long fixed forward rate position, is very thinly traded with variance ratios approaching unity, and daily microstructure effects from high-speed quoting (proxies by the wavelet correlations across the bid and ask) are indicated to be low. Second is an equity-like epoch, in which high-speed quoting is prevalent and whilst variance ratios rise they are never as high as those reported in Hasbrouck [2014]. Finally, the third epoch is very near delivery, with the future mimicking the implied short rate – volumes are high but quoting is less frequent; at this point the variance ratios at the shorter timescales sharply rise to levels much above those in the equity literature. I provide a comprehensive summary of these results and in particular I show that they are extremely stable across time and delivery quarter.

The remainder of this chapter is organized as follows: Section 2.1 provides a more detailed overview of the analysis of the microstructure of fixed income derivatives and summarizes related literature. Section 2.2 and Section 2.3 illustrate the empirical methodology and the unique Eurodollar dataset. Section 2.4 provides some analysis from the Bid-Offered volatility decomposition and Section 2.5 provides some summary comments.

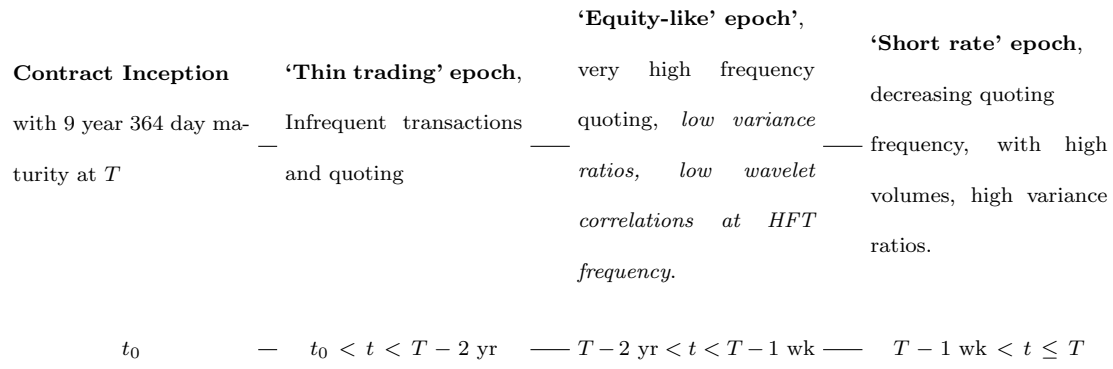


Figure 2.1: Trading 'Epochs' in the Life-cycle of a Eurodollar Futures Contract (yr denotes years and wk denoted weeks).

2.1 Background and Literature

The Eurodollar market, and specifically its microstructure, is not commonly studied directly in academic research papers, although the futures price is sometimes used as an ingredient in the construction of forward rate curves and as a forward looking asset pricing factor.[1] Given this lack of coverage in the literature I shall spend a short time reviewing the background of this market, the way the CME tapes are recorded and its general size characteristics; it is also useful to understand the function of the Eurodollar market in the international markets given the basic theoretical predictions on the characteristics of futures markets with differentiated agents. So historically who trades Eurodollar futures?

Eurodollar markets are defined as banking markets which involve Eurodollar short-term borrowing and lending conducted outside of the legal jurisdiction of

[1] In fact, Google Scholar and JSTOR record that no papers containing Eurodollar in the title or abstract have been published since 2010. An unbounded search shows that 488 mention Eurodollar in text with 12 directly related to pricing and none that address the transactional behavior of the market.

the United States authorities. Historically, the deposits have been mainly controlled by European banks and financial institutions, hence the name "European dollars". These banks were either foreign banks or foreign branches of the US banks. Because Eurodollars are not under the Fed regulation they offer higher margins than similar deposits within the United States.

The market developed after World War II following the Marshall Plan that resulted in large quantities of US dollars flowing into Europe and being deposited in European banks. Subsequent to the Soviet invasion of Hungary in 1956, the Soviet government was concerned that the US government could freeze its dollar deposit accounts in North American banks in retaliation. This led to the first Eurodollar account when the Soviet government transferred $800,000 into a British bank, which then deposited, anonymously, an equivalent amount into a US bank. This operation is considered to be the first time the words "European dollars" were used [Levi, 2007; Pilbeam, 2005]. The interest rate applied to the account is therefore the offered rate charged by the US bank to the British bank. Over time, this evolved into a standardized account and the most common duration is 3 months or 90 days pegged to the 3-month LIBOR rate. A natural evolution of this arrangement is to have a hedging instrument attached to forward delivery of a standardized account size. Hence, each Eurodollar future delivers one minus the prevailing dollar three-month LIBOR rate times one million dollars. The predominant literature on Eurodollar futures is on complete market pricing, see Brigo and Mercurio [2006] and Musiela and Rutkowski [2005]. Indeed, dollar interest rate hedging with Eurodollars had been a core feature of investment management courses in the UK and French institutions training eastern European central bankers since the late 1960s. However, this declined with the ICE-IR swaps boom of the early 2000s, even though for practical purposes much of the quoted swaps were either back-to-back or entirely synthetic with Eurodollar

strips.

The daily LIBOR has been the subject of extended discussion over the 2008 to 2014 period. On June 27, 2012, Barclays Bank settled with the US Commodities Futures Trading Commission, the US Department of Justice (DoJ) and the UK Financial Services Authority (FSA) fines totaling $200 million, $160 million and £59 million respectively. Up to this point, these were some of the largest penalties ever applied to a financial institution for manipulation of key rates. Subsequently, UBS, the Royal Bank of Scotland and J. P. Morgan have all seen fines in excess of this amount levied by these three regulatory authorities and the European Commission. In September 2013 ICAP, an electronic dealer broker also settled a lawsuit regarding rate manipulation of the LIBOR index paying $65 million to the CFTC and £14 million to the UK Financial Conduct Authority.

The objective of this study, however, is to analyze the market quote structure for interest rate derivatives; part of my focus will be on the structure of intraday variation around LIBOR announcements. Given the lack of knowledge of the exact form of the manipulation (in Section 2.2 I will outline some of the anecdotal evidence from the regulatory reports) it is almost impossible to pre-judge how the variation in bid and offered prices will change, and this study restricts itself to a "model-free" analysis of the market ex-post and concentrates on the transactional structure and in particular the impact of very high frequency transactions. However, from prior modeling approaches, such as the classic informed trading models of Kyle [1985], Glosten and Milgrom [1985] and Admati [1985], I can infer that the presence of informed traders should substantively increase prior to the updated news from the LIBOR fixing as a minority of traders exploit their private information.[1]

[1]The process of setting the LIBOR rate is called "fixing". This is not to be confused with cartel-like behavior to manipulate the market, so the expression "fixing the LIBOR" simply

The first LIBOR interest rate was published in 1986 [British Bankers' Association, 2008] and I shall now very briefly recap the reference rate-setting process. At 11.00 am each weekday morning the banks in the LIBOR pool submit their rates to an administrator. Up until September 2013, this was administered by the British Bankers' Association; this process is now administered by the NYSE: ICE LIBOR unit. Prior to September 2013, at 11.30 am Thomson Reuters would then publish the LIBOR fixings and make available the individual banks submissions. Since September 2013, the NYSE: ICE LIBOR unit makes the fixings available to various data vendors (including Thomson Reuters) for distribution to the market and directly to the futures exchanges for settlement on maturing interest rate derivatives set relative to this rate.

A key initial feature of the rate-fixing process was that banks "stated" their rates – no evidence of an actual transacted offered rate was needed. The original rationale for this was that banks may not have borrowed in all the maturities and currencies surveyed (e.g., there may not have been a 3-month deposit transaction for a particular bank). Therefore, if the rates needed to be backed by transactions then every so often more than 25% of the banks might provide a null rate and so no "mid-rate" might be computable. Given the importance of LIBOR in settling large volumes of derivatives and in the setting of borrowing rates on mortgages and corporate loans, the need for a complete set of offered rates appeared to outweigh the disadvantages of evidence of transactions behind the individual banks rate setting.

As such, unlike the Treasury bond future, the underlying instrument of the Eu-

means to set the rate. To actually force the rate to a particular level in order to try and make excess returns on interest rate derivatives, the most appropriate term is"manipulate the LIBOR fix" as used by the CFTC. The various settlement reports indicate two epochs for LIBOR rate manipulation: the first, pre-2007, was specifically to try and game money market derivatives trades and the second, a beauty contest, to present a picture of lower borrowing costs.

rodollar Future is not a bond, but a notional Eurodollar time deposits with interest rate pegged to the 3-month dollar LIBOR rate. The actuarial interest payment on a Eurodollar time deposit equals LIBOR times the numbers of days for the investment divided by 360 days [Brigo and Mercurio, 2006]. Let me set up some notational conventions to progress from the expressions for the theoretical model to the empirical model for the rest of the thesis.

I utilize two notational conventions for time denoted t. Subscript $t \in \{t, \ldots, T\}$ is a daily time index where $t \in \mathbb{N}_+$ is an integer. Alternatively, I set (t) (not subscripted) to be a continuous time index such that $t \in \mathbb{R}_+$ and more specifically I normally constrain $t \in [0, T]$ where $T = 1$ for one day. Later in the thesis, I will deal with an intraday transaction and quoting time, I will denote this by subscript $k \in \{1, \ldots, K\}$. Therefore, for every integer $k \in \mathbb{N}_+$ there is an equivalent calendar tick time entry $\tilde{K}_k \in \mathbb{R}_+$, such that $\tilde{K}_k + 1 - \tilde{K}_k \geq 0, \forall k \in \{1, \ldots, K\}$. In certain circumstances I can deal with identical time stamps; however, my general approach is to evenly distribute identical tick times over a single millisecond (the stated accuracy of the time stamp).[1]

A price quoted at time P_t is the end of day price, for all other measurements the stated value at time t is an average or a sum and this is declared locally. Alternatively, I use $P(t)$ to represent the price in continuous time, in general I think of a single indexed day as being in the domain $t \in [0, 1]$. I denote two arbitrary, but ordered dates as t_1 and t_2, such that $t_2 < t_1$ (normally $t_2 - t_1 = 90$ days for the three month LIBOR) and I utilize a time shift operator τ, such that for any date t I can define a time shift in days of $t + \tau$. Whilst unnecessary for my analysis I will follow the market convention and treat a year as having 360 and not 365 days.

[1] The data vendor informs me that the ordering is informative in terms of the receipt or execution of trades and quotes and I adjust the time stamps accordingly.

The intra-day index is more complicated as I deal with three clocks: the trade clock, the times when trades are executed; the inside quote clock, the times at which the best-bid and best-ask are updated (as a price-volume tuple); and the market depth clock, when any updates are made to the order-book at any level (reported as a price-volume-number of traders indexed by order book level tuple). Unfortunately, I have two uses of the word 'level'. First, I use it for the wavelet decomposition of the inside quotes; and second, it is used to describe the ranking, by price, of the quotes in the limit order book, both are entrenched terminologies and I shall be clear as to which context the work is used in throughout the thesis. Finally, the Eurodollars are quoted by contract maturity date and there are 40 concurrently traded quarterly dates, I therefore index by $i \in \{1, \ldots, 40\}$ and this is strictly in alphabetical order therefore $i = 1$ refers to EDH0 (the contract maturating on the third Wednesday of March 2000, 2010, 2020) and $i = 40$ is EDZ9 (the contract maturating on the third Wednesday of December 2009, 2019, 2029). Equilibrium pricing of a Eurodollar future is relatively well understood, the prices for a contract delivered at time t the price is $100 - \mathcal{L}(t, t+\tau)$, where $\mathcal{L}(t, t+\tau)$ is the quoted LIBOR rate from t to $t+\tau$.

Following from my notation set-up let the LIBOR at day t on a Eurodollar deposit with a maturity of τ days be expressed as

$$\mathcal{L}(t, t+\tau) = \frac{360}{\tau}\Big(\frac{1}{P^B(t, t+\tau)} - 1\Big), \tag{2.1}$$

where $\mathcal{L}$ is the LIBOR rate, P^B stands for a synthetic bond price from the implied delivery rate and $t \leq t+\tau$. I now consider a day T as a date for which all forward contracts are martingales. Assuming that t_1 and t_2 are future points in time and P^N is a notional amount, the basic payoff of a Eurodollar contract should be $P^N(1 - \mathcal{L}(t_1, t_2))$. As such, the fair price of Eurodollar future contract P_t^E can

be expressed as follows:

$$\begin{aligned} P_t^E &= \mathbb{E}_t\Big[P^N\big(1-\mathcal{L}(t_1,t_2)\big)\Big] \equiv P^N\big(1-\mathbb{E}_t[\mathcal{L}(t_1,t_2)]\big) \\ &= P^N\left(1-\mathbb{E}_t\Big[\frac{360}{t_2-t_1}\big(\frac{1}{P^B(t_1,t_2)}-1\big)\Big]\right) \equiv P^N\left(1-\mathbb{E}_t\Big[\frac{1}{4}\big(\frac{1}{P^B(t_1,t_2)}-1\big)\Big]\right) \\ &= P^N\left(1+\frac{1}{4}-\frac{1}{4}\mathbb{E}_t\Big[\frac{1}{P^B(t_1,t_2)}\Big]\right), \end{aligned} \tag{2.2}$$

where P^N is a notional amount and $\mathbb{E}_t$ is the expectation operator. For the 90-day Eurodollar contract referenced to the 3-month LIBOR, its future price $P^F(t,T)$ on the settlement day T is

$$P^F(t,T) = 100\big(1-\mathcal{L}(t,t+90)\big). \tag{2.3}$$

Therefore, the value of this Eurodollar Future is

$$\begin{aligned} P_T^E &= 100\big(1-\mathcal{L}(t,t+90)\big) \equiv 100\left(1-4\Big(\frac{1}{P^B(t,t+90)}-1\Big)\right) \\ &= 100\Big(1-\frac{1}{4}P^F(t,t+90)\Big). \end{aligned} \tag{2.4}$$

It is fairly evident that the future variation in $\mathcal{L}(\cdot)$ must be a semi-martingale ($\mathcal{SM}$) for there to be a "fair" valuation of the futures contract from the viewpoint of the long and short positions – see Brigo and Mercurio [2006] or Karatzas and Kardaras [2007] for an extended discussion of this requirement. Critically, as pointed out in Karatzas and Kardaras [2007], there need not be a complete absence of arbitrage, but the weaker form of 'No Unbounded Profit with Bounded Risk' for the construction of an appropriate numéraire for the price process. My analysis does not require any specific functional form; however, for the purposes of asset pricing my framework could be easily transposed from being a measure of market quality to a high frequency asset pricing factor.

That the fixing $\mathcal{L}(t, t+90)$ is a $\mathcal{SM}$ is a point of considerable recent discussion in the regulatory literature. As previously mentioned, up to 1 May 2014 Barclays Bank, UBS, Royal Bank of Scotland, Rabobank and Deutsche Bank have settled, or been fined for, allegations with regulatory bodies led by the US CFTC for manipulation of key rates. In the settlement reports the 3-month dollar LIBOR was noted several times as being a key target for manipulation of the fixing. [1]

2.1.1 Trading Eurodollar Futures

When designing statistical experiments on ultra-high-frequency data, understanding of the institutional arrangements within the transactional market matters. Eurodollar futures are traded in two ways, first via open-outcry and second via an electronic centrally cleared market. The most actively traded Eurodollar futures have quarterly deliveries and extend out to ten years from contract rollover to maturity. These contracts mature during the months of March, June, September, or December, extending outward 10 years into the future. Since the Eurodollar time deposits cannot be transferred or used as collateral for loans, the Eurodollar future cannot be made by physical delivery at maturity – settlement is always to cash. CME is the electronic and open-outcry clearing house for these contracts. The exchange operates continuous electronic trading from 5 pm (Chicago time) Sunday to 4 pm (Chicago time) Fridays. Open-outcry runs from 7:20 am (Chicago time) to 2 pm (Chicago time) Monday to Friday. The electronic trading market is the CME GLOBEX (centrally cleared limit order book) located centrally in Chicago. Unlike energy futures with physical delivery, no Eurodollars are cleared

[1] See http://www.cftc.gov/idc/groups/public/@lrenforcementactions/documents/legalpleading/enfbarclaysorder062712.pdf pp. 6–11 for the Barclays' Bank report. Page 6 specifically refers to "Three-Month Eurodollar Futures" and this forms the basis of the CFTC's remit for this section of the charge.

by over-the-counter (ClearPort) transactions, so my dataset is the complete history of the inside-quotes. One interesting issue is that of "iceberg" orders – these are permitted for this contract and are labelled by CME as a maxShow option. Cleared iceberg trades appear on the trades tapes as a cleared trade (often within the inside-spread) with volume attached (200 contracts is a very common iceberg trade size) much larger than the surrounding quotes. Whilst some anecdotal evidence on the price impact implication of iceberg trades exists within out dataset, I leave a comprehensive analysis of these trades for future work.

Most Eurodollar Futures are settled quarterly; the top code is ED and this covers both open-outcry and electronic settlement[1]. The codes for Eurodollars are denoted H (March), M (June), U (September) and Z (December). The last trading day is the second business day prior to the third Wednesday of the settlement month. Positions in excess of 10,000 contracts (recall that each contract is one million dollars) have limited trading (i.e. have position accountability) to the exchange and positions in excess of 850 contracts are reportable to the weekly CFTC commitment of traders report. As shown by the descriptives of my dataset, individual trades of in excess of 10,000 contracts are actually fairly common so the analysis of the Commitments of Traders (COT) report is relatively useful for this contract.

For each settlement quarter (H, M, U and Z), there are 10 maturities ranging from years ending in 0 to 9. Therefore 40 types of contract are simultaneously traded from EDH0 to EDZ9. The nearest maturity contract is the one delivering to coded quarter closest to the current date with a year ending in the same number as the present year. For instance, if today is March 11, 2014, then the nearest

[1]Note that Bloomberg tickers restrict the code ED to only electronically traded futures; my access to the historical CME tapes is by Thomson Reuters where GE reflects the electronic trades only.

delivery is EDH4 delivering on Wednesday, March 19, 2014. On Thursday, March 20, 2014, the EDH4 contract will then become the longest maturity contract, with settlement on Wednesday, March 20, 2024. The actual settlement is termed as being $T+2$, i.e. the rate is the rate on the last business day two days prior to the third Wednesday (this must be a business day in both New York and London, so very occasionally this moves). Rates are rounded to the nearest 1/100 of a basis point and are settled in cash. CME timestamps the Eurodollar futures prices using millisecond stamps and the tapes retain these tick-times. Table 2.1 illustrates the EDH0 quotes data over 4 min around noon GMT (6 am EST) on January 3, 2006 with the GLOBEX millisecond timestamps included. More recently a monthly contract for the "near years" has been issued and traded at a substantially lower volume than the nearest maturity quarterly contracts. My analysis indicates that the pattern for these contracts is consistent with the quarterly contracts in spite of the lower volume of trades – results can be found in my online appendix.[1]

In Figure 2.2 (a) and Figure 2.2 (b), I provide two daily snapshots of the best-bid best-offer, the mid-price and the executed trade prices for the March 2010 maturity Eurodollar future. Figure 2.2 (a) illustrates the trading in a 2.5-year maturity futures contract for September 10, 2007. We can see that the spread during the early part of the day is at nearly two basis points – very few transactions occur at this stage. As the day progresses, quote updates increase and so do the number of trades, some of which are outside the best-bid best-offer (the magenta trade markers). Quote updating at the inside-spread during this day is measured in seconds and even minutes. By contrast, Figure 2.2 (b) shows that when the same coded contract is observed within one year to maturity, the update speed declines

[1]The online appendix contains supplemental information about this thesis. As it is too large to upload to the submission system, you can find it at https://www.dropbox.com/s/uxbilz8879azdkb/Thesis_OnlineAppendix.pdf?dl=0

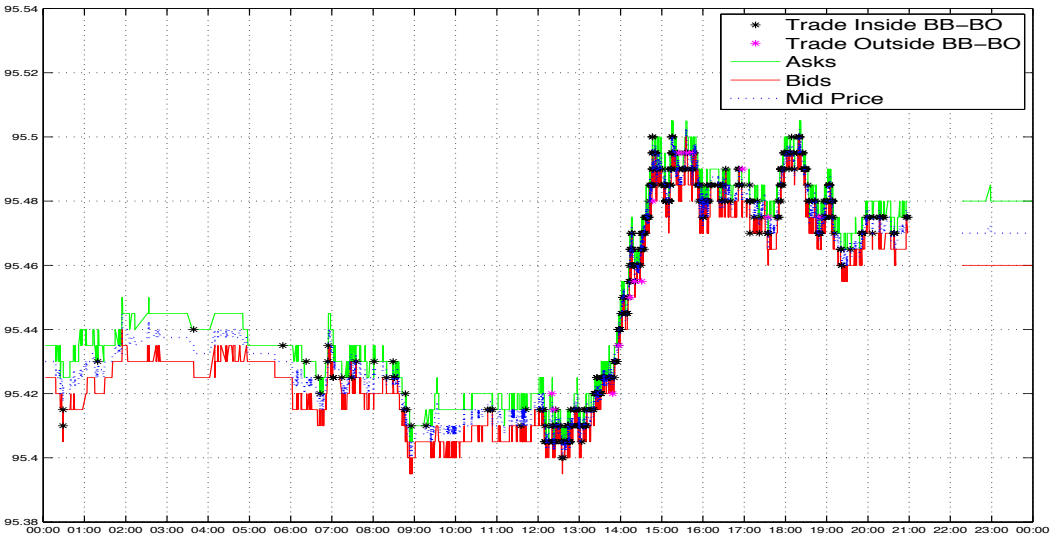
Table 2.1: Inside Quotes Data Structure

#RIC	Date[G]	Time[G]	GMT Offset	Type	Bid Price	Bid Size	Ask Price	Ask Size
EDH0	03-Jan-06	13:07:35.581	-6	Quote	95.03	50	95.065	50
EDH0	03-Jan-06	13:07:42.822	-6	Quote	95.03	50	95.07	50
EDH0	03-Jan-06	13:08:48.122	-6	Quote	94.715	1	0	0
EDH0	03-Jan-06	13:09:43.416	-6	Quote	94.995	2	0	0
EDH0	03-Jan-06	13:09:43.540	-6	Quote	95.025	50	95.09	2
EDH0	03-Jan-06	13:09:43.540	-6	Quote	95.025	50	95.065	50
EDH0	03-Jan-06	13:13:03.733	-6	Quote	94.715	1	95.065	50
EDH0	03-Jan-06	13:13:04.989	-6	Quote	94.99	2	0	0
EDH0	03-Jan-06	13:13:05.054	-6	Quote	95.02	50	0	0
EDH0	03-Jan-06	13:13:06.518	-6	Quote	95.02	50	95.06	50

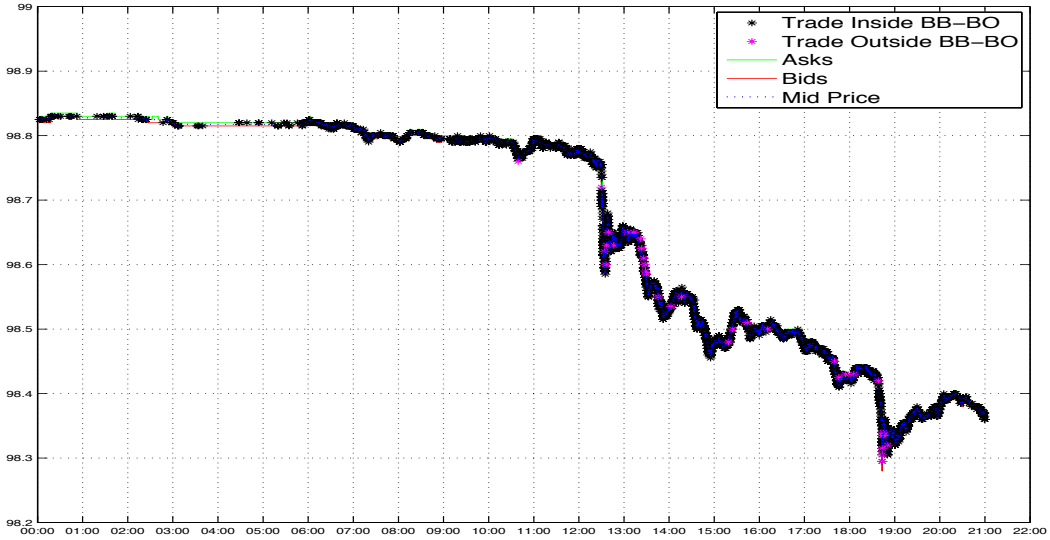
Notes: This table displays the raw inside quotes data structure of EDH0 quotes on January 3, 2006. I gather Eurodollar futures contract Quotes and Trades bid price and volume, ask price and volume, and trades price and volume in milliseconds. The data are sourced from the Thomson Reuters Tick History database (February 2014).

to sub-seconds and the number of transactions becomes higher. To my eyes, the spread does not vary much over the day, however we can see that quoting activity and density of transactions appear concomitant through the density of the trade markers. However, the actual variance and kurtosis of the spread over this day is far higher than is visible in the simple plot. There are a number of trades outside the best-bid best-offer, in Figure 2.2 (b) but the proportion is much lower than in the first trading epoch.

The extent of algorithmic trading is relatively unknown in this market. However, two facts support the conjecture that algorithmic trading of these contracts is both possible and present. The first rationale is the Globex market platform itself. The fastest access is via the GLink system which provides co-location options for traders and as such there is no technical reason to prevent high-frequency algorithmic trading. The second observation comes from the data. For contracts near maturity there is evidence of level of quoting and re-quoting at speeds beyond the reaction capacity of a human being (less than 200 milliseconds).



(a) Trades and inside quotes in EDH0 (March, 2010 maturity) on September 10, 2007



(b) Trades and inside quotes in EDH0 (March, 2010 Maturity) on June 5, 2009

Figure 2.2: Trades and Inside Quotes in EDH0 (March, 2010 maturity)

Notes: Figure (a) is trades and inside quotes in EDH0 on September 10, 2007: the best bid, best ask, the middle price and the executed price. The total volume on this day is $8.113 billion with 63,520 observations of asks and 63,534 observations of bids. Total trades observations are 1,016 with 14 trades outside the best bid and best ask. NOTE that the images are enlarged for review purposes only and are for illustrative purposes only. Figure (b) is trades and quotes in EDH0 on June 5, 2009, which has the largest daily trade observations for EDH0 with $449,988 billion volume of trades. Total trades observations are 30,030 with 34 trades outside the best bid and best ask. 435,611 asks and 435,544 bids are observed on this day. For data purposes I treat a day as being 00:00:00.001 CST to 11:59.59.999 CST, however, my plots follow the scientific literature and I plot the day from 00:00:00.001 UCT to 11:59.59.999 UCT.

I find a large number of reacting bids/asks that occur at "identical" millisecond stamps, indicating that the trade initiation is in fact faster than one millisecond. I will now review some of the current literature relating to algorithmic trading at high frequencies.

As a subset of algorithmic trading, high-frequency trading is usually defined as computer-based transactions that are designed and initiated at speeds faster than any human being could manage (sub-200 ms is a standard benchmark). The rise of HFT follows two developments: first, the relatively steep drop in the cost of high performance computing environments; and second, competition between exchanges resulting in increased speed of execution and far lower latency times.

2.1.2 A Brief Review of High Speed Quoting and Volatility

The ability to execute orders at millisecond and sub-millisecond speed has been the subject of considerable recent analysis, see Hendershott et al. [2011], Hendershott and Moulton [2011] and Riordan and Storkenmaier [2012] for a series of recent empirical work in this area. There is evidence that for some stocks, in particular large-cap stocks, automated trading has increased liquidity [Hendershott et al., 2011]. The increase in automated trading results in a reduction in bid-ask spreads, thereby reducing transactions costs for investors and indicating increased market depth. Similarly, Riordan and Storkenmaier [2012] report that after technological innovation resulting in latency reductions for HFT systems, liquidity increased across market capitalization. Hendershott and Moulton [2011] study the New York Stock Exchange (NYSE) system in 2006 and indicate that prices become more efficient due to faster price discovery – subsequently this

reduces the noise surrounding the efficient price of the asset.

HFT traders have been found to provide a high liquidity by keeping up prices and trading volumes at some level and smoothing the price volatility. Since HFT has been widely applied to exchanged trading, the most popular questions regarding HFT should be about the beneficial effects of HFT on the economy. Prior researches on this topic mainly examine the relationship between HFT and a variety of aspects of market quality, such as liquidity, volatility, depth, profitability and price discovery.

HFT as a mechanism for causing excess volatility in bid and ask quotes, resulting in excessive execution price risks for investors, is a well-studied conjecture. Since the flash crash in 2010, there has been general concern that volatility could increase and resiliency could degrade as technology improvements enhance trading speed. To discover how abnormal HFT activity varies with abnormal volatility, Brogaard [2012] adopts the NASDAQ-HFT and BATS-HFT datasets to estimate the short-term volatility. The findings show a positive relation between abnormal HFT marketable order participation and abnormal volatility.

Using the same NASDAQ-HFT dataset, Brogaard et al. [2014] investigate the related issue of the behavior of HFTs during times of market stress. The volumes of high-frequency traders increase on the daily highest volatility, which suggests that HFT is more active during market stress times. Brogaard et al. [2014] add a corollary to this analysis by suggesting that HFT actually improves price discovery on days with high intrinsic volatility. This 'critical-saturation' point for HFTs or ATs in the market is a finding I will demonstrate substantive evidence for in the latter part of my study. Using an alternative experimental setting, Hagströmer and Nordén [2013] assess the effect of passive HFT activities on short-term volatility based on tick size changes. The findings indicate that increased

passive, market-making, activities have the ability to cause a decrease in short-term volatility, although the density of HFTs is not studies in this instance.

At a structural level, away from the technological evolution, the relation between the price volatility and futures contracts with respect to time to expiration has been comprehensively studied. The classic theoretical treatment being Samuelson [1965] who proposes that futures price volatility should increase when contracts are close to expiration. This behavior is denoted as the 'maturity effect' in the extant literature. A great number of empirical studies, using a variety of experimental approaches, have been based around detecting evidence for or against the Samuelson conjecture; albeit with very mixed results see [Anderson and Danthine, 1983; Bessembinder et al., 1996; Bollen and Inder, 2002; Chen et al., 1999; Daal et al., 2006; Galloway and Kolb, 1996; Grammatikos and Saunders, 1986; Gurrola and Herrerías, 2011; Han et al., 1999; Kalev and Duong, 2008; Khoury and Yourougou, 1993; Milonas, 1986] for about a 50:50 split on confirming the volatility adjustment versus rejecting it.

In my study I will combine the HFT literature with the analysis of maturity effects in futures contracts to build a picture of the microstructure lifecycle of a given contract. Maturity effects have usually been studied in relation to agricultural futures contracts rather than financial futures [Bessembinder et al., 1996; Kalev and Duong, 2008], and historically the focus has been on trying to explain why the maturity effect is not working for all markets from an economic perspective. Studies such as by Bessembinder et al. [1996] point out that if a market has negative covariance between spot price changes and change of net carry costs, this market usually appears to have the maturity effects. Bessembinder et al. [1996] also indicate that negative correlation and covariation usually occur in physical assets instead of financial assets. In contrast, Hong [2000] argues that the

variation of informed trading in the market have different effects on the maturity effects of return volatility as contracts roll to the expiration. When the degree of information asymmetry is relative low, the Samuelson effect exists in the future market. While the Samuelson effect does not hold in the future market when the degree of information asymmetry is relative high. Comparative studies such as Kalev and Duong [2008] confirm that the maturity effect usually follows with the negative covariance condition, and they find little support for visible maturity effects in the Eurodollar futures market. This is in contrast to my findings, which indicates that the maturity effect in Eurodollar futures is primarily in the variance ratios as the contract approaches maturity and that the causation for this change in the variance scaling varies as I see high frequency trading activities peak and then dip just prior to maturity, with one year to two months from maturity giving the highest proportion of high frequency, possibly speculative, trading.

The most closely related prior study to this chapter is by Hasbrouck [2014]. To evaluate the short-term quote volatility in the US equity market, Hasbrouck examines the variance ratios at different timescales from 50 ms to 27 min for the US equity National Best Bid and Offer (NBBO). An important contribution of Hasbrouck's study is in analyzing volatility in prices, rather than in returns. Inside-quote volatility at high frequency is determined over 16 wavelet levels using both short-term data in 2011 and long-term historical data from 2001 to 2011. At short-term quote volatility in the US equity market, Hasbrouck [2014] demonstrates that short-term quote variance is larger than variance in long timestamps. Moreover, bid-ask correlations are positive but low at sub-second timescales, which suggests that the bid and offer are not moving together. Hasbrouck utilizes a simulation approach to estimate millisecond-level volatility analysis with second-stamped quotes data. He finds high correlations between the original timestamp estimation and the simulated estimation. For the long-term

historical analysis, the highest quote volatilities occurred during the 2004–2006 period, this coincides with the transition to electronic trading in the US markets and a share rise in HFT.

2.2 The Hasbrouck Wavelet Approach

Hasbrouck [2014] employs a Haar wavelet method to evaluate the information contained in high-frequency data. Wavelet analysis has been widely used for processing fundamental signal tasks, such as compression, abatement of noise, or optimizing a recorded sound or image. Wavelets are widely utilized in economics and finance as well, for a non-exhaustive list of applications see: [Conlon and Cotter, 2012; Conlon et al., 2008; Galagedera and Maharaj, 2008; Gencay et al., 2002; Lien and Shrestha, 2007; Percival and Mofjeld, 1997; Percival and Walden, 2000; Rua and Nunes, 2009]. As with all wavelets, the Haar basis is a multi-resolution analysis; that is the decomposition of a signal into sub-signals of different size resolution levels, the number of partitions is given by 2^{levels} as the sample is progressively partitioned. Whilst my approach is effectively in line with Hasbrouck [2014], some differences are present given the varying nature of my data, hence I review the implementation in some detail.

The wavelet transform is often compared with the Fourier transform in so far as: the raw format of most signals is in the time-domain; the Fourier transform determines the frequency content of signals; the transform is invertible and both continuous and discrete time approximations exist. However, there are limitations to the Fourier transform often not present in a wavelet. Time information is lost after transforming the signal to the frequency domain, so that it becomes impossible to tell what happened in some periods. Many signals contain inter-

esting transient response characteristics, such as drifts, trends, sudden changes, even the beginning and end of signals. These features have all been shown to be characteristics of financial data, especially high and ultra-high frequency data.

To overcome the shortcomings of the Fourier approach, Gabor [1946] proposes a short-time Fourier transform (also known as the Gabor transform). The short-time Fourier transform is a technique to transform a time-domain signal to two-dimensional functions with both the time and frequency domains. It provides the information contained in signals in a certain time and a certain frequency range. The accuracy of the information depends on the time window size. However, time window sizes of the short-time Fourier transform are the same for all frequency components. Many signals need diverse time windows to obtain more accurate time and frequency information.

The wavelet transform is an advanced approach to overcome the limitations of both the Fourier transform and Gabor transform to obtain further information. Generally speaking, the main difference between the wavelet transform and the Fourier transform is that wavelets can be localized in both time and frequency, while the Fourier transform is only localized in frequency. Since the wavelet transform combines information from the time domain and the frequency domain, it has more flexibility than the Fourier transform. Compared with the short-time Fourier transform, the wavelet transform proposes time windows of flexible width. When low-frequency information is needed, the wavelet transform uses a long time window. A short time window can be employed to accurately analyze high-frequency information.

The Haar wavelet is an example of an orthogonal system of functions defined on the unit interval [0,1] [Haar, 1910]. There are two functions that play key roles in wavelet analysis: the scaling function (father wavelet) and the wavelet

(mother wavelet). The scaling function $\mathscr{S}(k)$ corresponding to the Haar wavelet can be defined by $\mathscr{S}(k) = 1|0 \leq 1$ or $\mathscr{S}(k) = 0|\forall k > 1$. The Haar wavelet's mother function $\mathscr{H}(k)$ is defined over the domain $\mathscr{H}(k) = 1|0 \leq k < 1/2$ or $\mathscr{H}(k) = -1|1/2 \leq k < 1$ and 0 otherwise. The scaling equation of the Haar wavelet is therefore: $\mathscr{S}(k) = \mathscr{S}(2k) + \mathscr{S}(2k+1)$ and hence $\mathscr{S}(k) = \frac{1}{\sqrt{2}}\sqrt{2}\mathscr{S}(2k)+\frac{1}{\sqrt{2}}\sqrt{2}\mathscr{S}(2k-1)$, so the Haar wavelet filter coefficients are $\gamma_0 = \gamma_1 = \frac{1}{\sqrt{2}}$. The Haar mother wavelet contains a short positive pulse followed by a short negative pulse, as such the Haar wavelet belongs to an orthogonal system in the field of real numbers. In a useful result for asset pricing Paul Levy demonstrated that the Haar wavelet is more accurate in its local approximation if subdivided into different intervals, while the Fourier functions only have one interval Meyer [1993] provides a comprehensive summary of the statistical properties of the Haar wavelet amongst others.

My objective is to compute the ratios of the bid (subscript B) and ask variance (subscript A) and covariance (subscript AB) at differing timescales, usually relative to the daily variation. Let the bid, ask and trade prices of Eurodollar Futures be considered as a discrete sequence of prices $P = \{P_{i,1}, P_{i,2}, \ldots, P_{i,t}\}$. The variance and covariance ratios are denoted respectively $\tilde{\mathscr{R}}_B$, $\tilde{\mathscr{R}}_A$ and $\tilde{\mathscr{R}}_{A,B}$. In each case, I have a variance–covariance ratio operator of the form:

$$\tilde{\mathscr{R}}_i[v, v'] = \begin{cases} \frac{1}{\mathbb{M}_v}\sigma^2_{i,\mathfrak{T}_v}.(\frac{1}{\mathbb{M}_{v'}}\sigma^2_{i,\mathfrak{T}_{v'}})^{-1}, & i \in \{A, B\} \\ \frac{1}{\mathbb{M}_v}\sigma_{i,\mathfrak{T}_v}.(\frac{1}{\mathbb{M}_{v'}}\sigma_{i,\mathfrak{T}_{v'}})^{-1}, & i \in \{AB\} \end{cases} \tag{2.5}$$

setting $d \in \{v, v'\}$ to represent to time scale indices v and v',

$$\sigma^2_{i,\mathfrak{T}_d} = \mathbb{E}[(P_i - \mathbb{E}[P_i])^2]|\mathfrak{T}_d, \quad i \in \{A, B\} \tag{2.6}$$

$$\sigma_{i,\mathfrak{T}_d} = \mathbb{E}[(P_A - \mathbb{E}[P_A])(P_B - \mathbb{E}[P_B])]|\mathfrak{T}_d, \quad i \in \{AB\} \tag{2.7}$$

where $\mathcal{T}_d$ represents an exact time scaling (milliseconds, seconds, minutes, hours, days), $\sigma^2_{i,\mathcal{T}_d}$ for $i \in \{A, B\}$ is the variance at that time scale with $\sigma^2_{i,\mathcal{T}_d}$ for $i \in \{AB\}$ is the covariance, and $\frac{1}{\mathcal{M}_d}$ $j \in \{v, w\}$ is a factor that coverts $\sigma^2_{i,\mathcal{T}_d}$ and $\sigma^2_{i,\mathcal{T}_d}$ into a common time scale. If prices follow a random walk with constant volatility then the ratio operator $\tilde{\mathscr{R}}_{i,d} \to 1$. Although it should be noted that this is not a sufficient condition to identify the time series as a random walk. The conditional suffix $|\mathcal{T}_d$ denotes the time scaling that the expectation $\mathbb{E}[.]$ operates over. It is useful to have a general labelling of the time scale due to the properties of the filter that I will apply to estimate the scaled average price $\mathbb{E}[P_i]$.

The focus of my analysis is on the bids and asks components and computing $\sigma^2_{i,\mathcal{T}_v}$ and $\sigma_{i,\mathcal{T}_v}$. Notice that the variances are in terms of price levels and not returns, Hasbrouck [2014] offers a quite convincing argument that when estimating microstructure effects the deviations from the efficient price will, in all likelihood, be transient, rather than persistent.

I use a wavelet filter to a) identify the time scale and b) provide a local average price. I can think of the time scaling as being a series of local averages and the wavelet provides a convenient systematic approach to determining the local time scales. The basic wavelet constructs are the basis functions and these are formed according to the rule $\mathscr{H}_w(k) = 2^{w/2}\mathscr{H}(2^w k - d)$, where $w \in \mathbb{Z}$ denotes as a unique time scale level, $d \in \mathbb{Z}$, which sets out the scheme for determining the series of local averages at scale d.

One of the fundamental advantages of wavelet analysis is the capability to decompose time series into different components. Based on the concept of multi-resolution analysis, a time series can be decomposed into different elements associated with a unique timescale level w. Figure 2.3 represents the multi-level wavelet decomposition process. The original signal is split into high-frequency

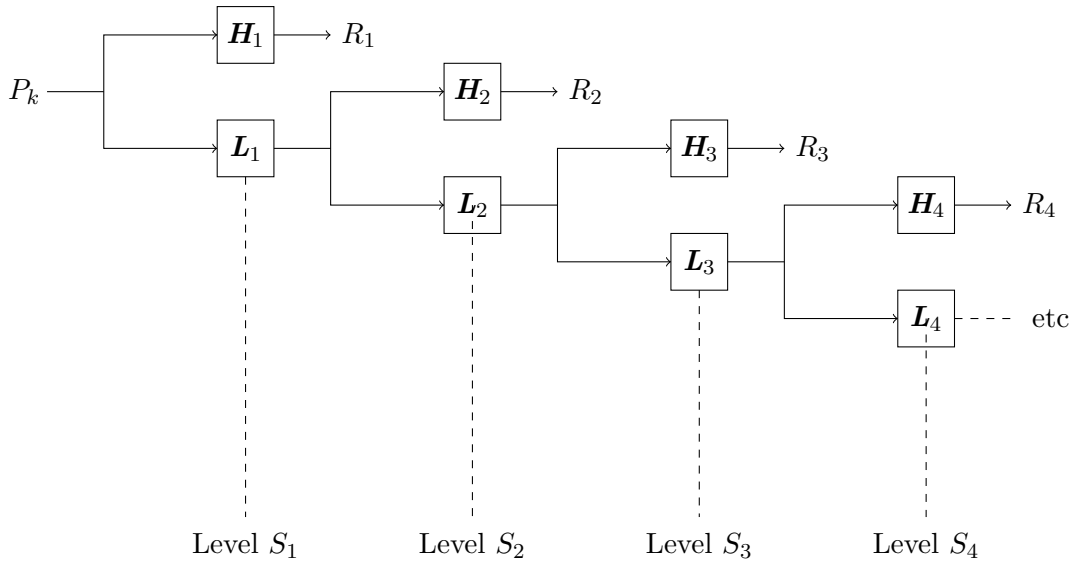


Figure 2.3: Wavelet Decomposition Tree

Notes: This figure represents the multi-level wavelet decomposition process. The original prices (P_k) is split into high-frequency components ($\boldsymbol{H}_w$) and low-frequency components ($\boldsymbol{L}_w$), where k is the intraday time index and w is the timescale levels. For a length 2^w vector of observations, the vector S_w is a smooth component, which equals to the sample mean and defined the w^{th}-level wavelet smooth. The residual R_w stands for a rough component at timescale level w.

components ($\boldsymbol{H}_w$) and low-frequency components ($\boldsymbol{L}_w$). For a length $l = 2^w$ vector of observations, the vector S_w is equal to the sample mean and defined the w^{th}-level wavelet smooth, while R_w stands for the residuals. I can think of these as being 2^w local averages. Hasbrouck [2014] argument is that the microstructure noise components are stationary in levels (as opposed to returns for GARCH type models) and that a local averaging approach provides a better approximation of the central tendency of the variance than differencing. Indeed, as is demonstrated in Meyer [1993], locally flat domains are easily filtered by the wavelet and subtracted from cross wavelet covariance estimation, a feature that causes standard quadratic regression methods considerable difficulty. Notation-

ally, therefore, wavelet multi-resolution analysis can be defined as:

$$P_w = S_w + R_w \tag{2.8}$$

I assume that the wavelet smooth S is the local mean of length l ending at a time k, which can be calculated as $S(l,k) = l^{-1}(p_k + p_{k-1} + \cdots + p_{k-l-1})$ where the deviation equals $R(l,k,\tilde{s}) = p_{\tilde{s}} - S(l,k)$ and $\tilde{s}$ is a tick-time, such that $k - l < \tilde{s} \leq k$.

Eurodollar Futures prices $P_k = P_{k-1} + \varepsilon_k$, where ε_k is a white-noise process with unit variance. Hence, the mean square deviation, $\mathfrak{M}$, can be expressed as

$$\mathfrak{M}(l,k) = l^{-1} \sum_{\tilde{s}=k-l+1}^{k} [R(l,k,\tilde{s})]^2 \tag{2.9}$$

The expectation can be simplified to $\mathbb{E}[\mathfrak{M}(l,k)] = Var[R(l,k,\tilde{s})] = \sigma_l^2$. The mean $S(l,k)$ is a smooth component, the residual series $R(l,k,\tilde{s})$ is a rough component. So the rough variance is $\sigma_l^2 \equiv \sigma_{lw}^2$ at the timescale $l_w = l_0 2^{w-1}$, where σ_l^2 reflects variation at timescale l_w and shorter. The incremental change in moving from timescale level $w-1$ to w is $\nu_w^2 = \sigma_w^2 - \sigma_{w-1}^2$, where ν_w^2 reflects the incremental variance (or wavelet variance) at timescale level w and σ_w^2 is a rough variance. Hence I can approximate the exact time scaled $\tilde{\mathscr{R}}_{i,d}$, from the nearest $\tilde{\mathscr{R}}_{i,w}$ computed from the scaled ratio of σ_w^2 to another scale $\sigma_{w'}^2$.

Interpretation of the covariance ratio is relatively clear, if the covariance decreases (that is the covariation at high frequency is less than the covariation at the lower frequency) then the bid and ask prices have decoupled. Some degree of decoupling is to be expected as the update of order flow is discrete at the highest frequency, hence one price (the bid or the ask) will be expected to lead the other at different

points. Over a day this will even itself out, so as the frequency decreases, the degree of coupling will naturally increase.

If high frequency trading increases liquidity and decreases latency, then the degree of decoupling should be less when there are more HFTs in the market. However, if a HFT from ATs is designed to deliberately obfuscate the order-flow to enable a manipulative market trading strategy, then I might anticipate that the degree of decoupling will be higher and hence $\tilde{\mathscr{R}}_{AB}[w, w']$ will decrease. In terms of variance the ratios $\tilde{\mathscr{R}}_{A}[w, w']$ and $\tilde{\mathscr{R}}_{B}[w, w']$ denote the excess volatility at wavelet time scale w relative to w'. Hence, higher levels of $\tilde{\mathscr{R}}_{A}[w, w']$ and $\tilde{\mathscr{R}}_{B}[w, w']$ represent a higher degree of risk for buyers and sellers at the point of execution.

2.3 Data and Summary Statistics

In this section I introduce the data processing procedures and provide a summary of the descriptive statistics of the sample. I have constructed a unique dataset for the 40 quarterly Eurodollar Futures contracts from January 1, 1996 to December 31, 2013. My futures data was obtained from CME exchange tapes via Thomson Reuters and is available for researchers to purchase, given the volume of data within the order book and the long sample of time this is likely to be the largest study of its type ever conducted. Each is labelled by a Reuters Instrument Code (RIC), such as EDH0, EDH1, ..., EDZ9 (see Table 2.2). The 40 Eurodollar Futures contain all Eurodollar quotes and transactions on the Chicago Mercantile Exchange's Globex platform. The last trading day for Eurodollar contracts is two business days prior to the third Wednesday of the delivery month. Table 2.2 shows the settlement date, duration, the average of bid price, the average of bid volume, the average of ask price, the average of ask volume, the average of trade price

and the average of trade volume for ED contracts having their delivery months in the same year.

Table 2.2: Size Data Sample for Time and Sale Data

RIC	Duration	Trades		Asks		Bids	
		Average Volume ($ trillion)	Average no. of obs. (million)	Average Volume ($ quadrillion)	Average no. of obs. (million)	Average Volume ($ quadrillion)	Average no. of obs. (million)
ED?0	2000-2010	28.22	1.64	47.38	23.04	46.09	22.92
ED?1	2001-2011	15.00	1.50	74.23	23.34	74.73	23.39
ED?2	2002-2012	12.54	1.45	69.64	24.05	68.23	24.20
ED?3	2003-2013	12.69	1.38	181.74	24.05	183.37	24.17
ED?4	2004-2014	22.11	1.62	171.92	24.53	166.41	24.60
ED?5	1995-2005	38.99	1.92	83.23	25.65	80.73	25.79
ED?6	1996-2006	43.77	1.68	56.75	25.17	57.18	25.23
ED?7	1997-2007	58.98	1.36	65.88	21.88	66.62	21.98
ED?8	1998-2008	69.24	1.62	35.92	21.78	36.49	21.92
ED?9	1999-2009	49.23	1.90	19.85	23.31	19.74	23.32

Notes: This table depicts data sample for time and sale data. The roll-over dates of the 40 Eurodollar quarterly trade futures and the average number of observations plus the average volume of bids, asks and trades. The time and sale data are from the Thomson Reuters Tick History database for January 1, 1996 to January 1, 2014, labelled by the Reuters Instrument Code (RIC; EDH0, EDH1, ..., EDZ9). I calculate the average value for ED contracts having their maturity date in the same year, denoted ED?0, ED?1, ..., ED?9. Hence, ED?0 includes four 10-year future contracts from 2000 to 2010 – namely, EDH0, EDM0, EDU0 and EDZ0. Note, that the minimum tick size on the exchange is 1/4 of a basis point for the nearest expiring contract and 1/2 otherwise.

Specifically, ED?0 includes four 10-year future contracts from 2000 to 2010, namely, EDH0, EDM0, EDU0 and EDZ0 (see Table 2.2). Their delivery months are March 2010, June 2010, September 2010 and December 2012, respectively. The average number of ED?0 bids and asks are both 23 million approximately and the number of trades is 1.64 million. The average volume of ED?0 trades, asks and bids are $28.22 trillion, $47.38 quadrillion and $46.09 quadrillion, respectively. Across the 40 ED contracts by year of delivery, ED?3 has the largest

average bids volume ($183.37 quadrillion) and the largest average asks volume ($181.74 quadrillion). The largest trades volume is the ED?8 contract with $69.24 trillion. Furthermore, ED?5 has the largest average number of bids, asks and trades, with approximately 25.79 million observations of asks, 25.65 million observations of bids and 1.92 million observations of trades.

To avoid unnecessary replication, I report the EDH? (? $\in \{0, \ldots, 9\}$) futures only in this chapter. However, my results are consistent across the contracts, and these results are available in the online appendix which provides a variety of subsampling and robustness checks.

Figure 2.4 plots the best bid, best ask and trades for the EDH0 from 1996 to 2013. EDH0 is a 10-year future contract (settlement at March 2000, 2010 and 2020) with the most recent settlement date being March 17, 2010. The number of updated bids and asks are both 23 million approximately, and the updated ask and bid volumes are $35 quadrillion and $38 quadrillion respectively. Bids prices, asks prices and trades prices of EDH0 are marked by the red line, green line and blue line, respectively, although the trades are masked almost entirely by the roughly 20:1 quote to trade ratio.

The long-term evolution of EDH0 illustrated in Figure 2.4, it plots the evolution of this contract from long forward rate to near-spot forward rate over these three epochs. The most dramatic shift is the near collapse in the short rate during 2008. Over this period, some 24 million quotes were submitted, with the notional value of outstanding interest in the quadrillions of dollars. After gathering all 40 Eurodollar Futures data I split them into two parts. The first sub-sample is from January 1, 1996 to July 31, 2007; the second is from August 1, 2007 to December 31, 2013. Then I eliminate all the data with prices smaller than $80. For the remainder of the chapter, I will use days to maturity as the time index, instead

of the actual calendar date.

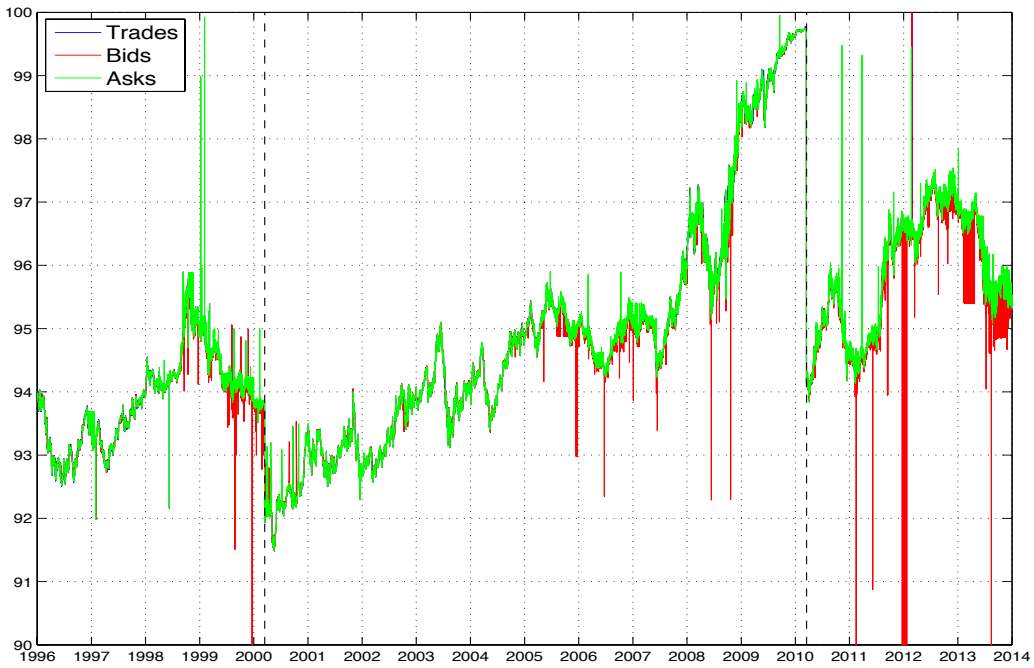


Figure 2.4: EDH0 Contract Inside Spreads

Notes: This plot depicts EDH0 inside quotes from 1996 to 2013, for 2000, 2010 and 2020 deliveries. The two rollovers of this contract code are denoted by the dotted lines.

Table 2.3 reports the descriptive statistics for each EDH future contract in these two periods. Descriptive statistics of the other 30 ED contracts can be found in the online appendix. Mean, median and mode of bid-ask spread are in basis points (bps) and mean returns are in one-millionth. For the first sub-sample (see Table 2.3 period A), the highest mean bid-ask spread belongs to EDH2 contract with 2.22 bps, but its median spread is only 0.5353 bps. The mean and median are far apart for EDH2, with 805.8972 in the kurtosis of the spread. The highest median spread occurs in EDH1 contract with 2.1042 bps, which also has the highest mean trades return. The smallest mean bid-ask spread belongs to EDH6 – only 0.6091 bps. Although EDH7 has the highest bids volume with $63.842

quadrillion, its mean trades return is the lowest among all the ten EDH futures.

In the second sub-sample (see Table 2.3 period B), EDH7 has both the highest mean spread and mean returns. Due to the extreme value existence, its median spread is only 0.5149 bps with 2,337 kurtosis spread. EDH7 also shows the highest mean returns. On the other hand, EDH0 expresses the lowest mean spread and mean return. Furthermore, EDH4 has the highest asks volume with $220.811 quadrillion. The largest trade volume is the EDH9 contract with $52.2 trillion. Meanwhile, EDH3 has the largest asks observations with approximately 24 million. The differences are relatively minor compared to the overall quantity of trading, for contracts maturing in 2009 to 2011 trading volume is high and this indicates that the safety of Eurodollar deposits were in demand at this time of financial market stress. It is noticeable that the spreads widened during this period and the price of these contracts fell, this can be seen in Figure 2.4.

More specifically, the mean bid-ask spread of EDH0 in the second period is smaller than that before 2007. The same pattern appears in the median and mode of its bid-ask spread. For EDH0 trades prices, we can see its mean returns, maximum returns and minimum returns. However, the maximum bid-ask spread, skewness spread and kurtosis spread after 2007 are larger than in the first period. This indicates that the closer to the maturity date, the narrower the bid-ask spreads and the higher the frequency of trading for ED contracts. What is particularly interesting is that the contracts maturing throughout the 2007–2012 period have the highest kurtosis indicating that even for a market with the renowned depth of the Eurodollar, substantial liquidity shocks were felt.

Table 2.3: First Quarter Delivery Contracts Statistic Description

RIC	Mean spread (bps)	Median spread (bps)	Mode spread (bps)	Max spread	St.dev spread (bps)	Skewness spread	Kurtosis spread	Mean returns (one millionth)	Max returns	Min returns	St.dev returns (bps)	Skewness returns	Kurtosis returns
Period A: Januray 1996 - July 2007													
EDH0	1.4575	1.0567	1.0510	0.0547	3.2725	42.3113	2,717.6928	0.0943	0.0138	-0.0198	2.0361	-9.6647	2,681.4828
EDH1	2.1979	2.1042	0.0000	0.0525	4.3642	27.4068	1,242.8045	0.0994	0.0108	-0.0201	1.7563	-17.4171	3,318.1755
EDH2	2.2185	0.5353	0.0000	0.0726	8.4861	23.6435	805.8972	0.0811	0.0109	-0.0563	2.2065	-165.9489	42,328.8565
EDH3	2.1448	0.5242	0.0000	0.0425	7.4073	14.8266	333.8133	0.0595	0.0092	-0.0556	1.8320	-199.0733	60,534.4976
EDH4	1.6973	1.0106	0.5059	0.1115	4.7529	37.7851	6,385.3045	0.0616	0.0081	-0.0623	1.9863	-224.2629	70,600.9591
EDH5	0.9052	0.5142	0.5153	2.2516	27.9470	798.4578	643,162.1358	0.0162	0.0279	-0.0302	0.9041	-74.3419	60,354.1097
EDH6	0.6091	0.5216	0.5250	0.0579	1.0275	83.0071	18,832.3889	-0.0123	0.0400	-0.0405	0.8555	-36.9462	133,763.0831
EDH7	0.6685	0.5271	0.5284	0.0620	2.7869	67.0657	6,077.5817	-0.0166	0.0043	-0.0260	0.6268	-101.5777	39,998.7236
EDH8	0.6715	0.5267	0.5258	0.0620	1.1519	134.6662	41,984.5674	0.0094	0.0059	-0.0097	0.5780	-8.6882	2,470.0066
EDH9	0.8650	0.5277	0.5246	0.1156	1.8810	152.5455	68,286.2488	0.0389	0.0442	-0.0442	1.5691	-7.0558	51,223.3289
Period B: August 2007 - December2013													
EDH0	0.9609	0.5170	0.5014	0.1553	1.9434	131.6144	70,698.0733	0.0033	0.0079	-0.0567	0.5704	-609.5429	607,831.4673
EDH1	1.3179	0.5118	0.5021	0.1848	3.9274	87.3828	27,614.9045	0.0044	0.0076	-0.0533	0.5814	-535.1033	493,029.1265
EDH2	1.3079	0.5138	0.5026	0.1830	3.7880	93.1862	29,520.3413	0.0047	0.0033	-0.0363	0.4512	-376.3398	302,626.1256
EDH3	1.6970	0.5112	0.5015	0.0559	3.7869	9.5710	295.9432	0.0055	0.0368	-0.0368	0.6860	-121.5284	205,078.3878
EDH4	1.4035	0.5098	0.5021	0.1769	6.5587	38.2543	4,722.5535	0.0422	0.0232	-0.0231	0.9732	0.3991	4,965.5213
EDH5	1.5816	0.5080	0.5031	0.4767	16.1712	136.3680	30,338.9690	0.0451	0.0055	-0.0046	0.6486	0.4279	296.7271
EDH6	1.3777	0.5101	0.5053	0.0420	9.1739	19.5081	409.5095	0.0481	0.0050	-0.0066	0.8672	0.2047	205.7452
EDH7	1.9253	0.5149	0.5088	0.1839	12.8469	30.6996	2,336.8586	0.0625	0.0081	-0.0072	1.0535	0.1052	256.5061
EDH8	1.8702	0.5246	0.5230	0.1876	10.6464	29.6438	2,167.1580	0.0163	0.0053	-0.0308	0.9274	-44.8187	15,030.5633
EDH9	1.1304	0.5188	0.5159	0.1553	9.7657	64.4546	4,384.0072	0.0037	0.0067	-0.0313	0.4439	-205.5758	147,461.4013

Notes: This table illustrates the first quarter delivery contracts descriptive statistics between 1996 and 2013 with maturity in March. The mean, median and mode of the bid-ask spread are presented in basis points and mean returns reported in one millionth in contrast to the yearly deliveries I now cut one quarterly delivery date (H for March) by year of delivery.

2.4 Term Structure Properties of the Estimated Quotes Volatility

The CME tapes are in the form of large comma separate variable (csv) files containing monthly blocks of milliseconds timestamped inside-bid and inside-ask quotes and volumes by contract type (ED{quarter}{delivery year last digit}). I convert the data in the original files into hdf5 format and then deploy the wavelet approach of Hasbrouck [2014] described previously for the 40 ED contracts.[1] Many studies of high-frequency data remove outliers and other possibly erroneous records. The most complex aspect of my data is the variation between open-outcry and electronic quote updates. However, I have access to the GE-coded data tapes containing only the electronic records and I find no discernible difference in the results. It should also be noted that the CME open-outcry quotes are three orders of magnitude less in numbers of updates on average across contracts. Analysis of the CME open-outcry pit trades themselves offers some variation. However, the small number of completions makes this less economically meaningful as an object of study.

The wavelet approach requires some choices on behalf of the econometrician – most of these involve the number of levels to filter, plus the choice of basis. Hasbrouck [2014] utilizes 16 levels for equity data for his highest frequency tranche of TAQ data and a Haar wavelet basis (local average filter). However, for this study I have adopted nine levels, plus the zero level ($J = 10$) and maintained the wavelet basis. This is because Eurodollars futures quotes tend to update at

[1]The Matlab codes used to calculate the bid and ask wavelet variances are adapted from those provided on the webpage of Joel Hasbrouck and the authors gratefully acknowledge their use. I use a hierarchical data format, hdf5, to the streaming of daily data from the tapes to the wavelet algorithm and reduces the memory problems associated with such large datasets over relatively long time periods.

slightly lower frequencies than equities for the majority of my sample (although at $1 million per contract the dollar volume per quote update is far larger). My timescales reveal a decline in average tick update time for my selected treatment groups from 30 days before maturity to 1 year before maturity. For example, level 1 timescales of EDH0 bid wavelets in the pre-30-day group, pre-60-day group, pre-3-month group and pre-1-year group are 8.9, 4.3, 2.5 and 1.3 seconds, respectively. The longest timescales, level 9, in these four pre-maturity groups are 38.2, 18.5, 10.9 and 5.7 minutes. Recall that Eurodollar Futures trade 24 hours continuously during CME opening times so there is less compression of trades into fixed exchange hours as occurs in the case of the equity and equity options market. In the online appendix, I present the median for all timescales of bids wavelet variance, asks wavelet variance, bid–ask wavelet covariance and wavelet bid–ask correlations for EDH0 to EDZ9. A series of other volatility results for these 40 Eurodollar contracts are reported in the online appendix, such as the bootstrap mean, mean, minimum and maximum and standard deviation.

After the settlement of the shortest maturity contract the exchange rolls over the contract code and hence the maturity of the contract extends out to 9 years and 362 days. At this point, the level of quoting and trading activity drops towards zero. For days with more than 5,000 quotes, daily bids and asks wavelet variances and bid/ask wavelet covariance have been computed. I have then cut the results of this daily analysis and computed average bid and ask variance and bid–ask covariance for 30 days, 60 days, 3 months, and 1 year before ED contract maturity date. For these four sub-samples, I calculate all the timescales for nine levels for EDH contracts. Within the sub-groups, the wavelet results are classified based on prices. So I have different timescales for 0–25% prices, 26–50% prices, 51–75% prices and 76–100% prices in the 30-day group, 60-day group, 3-month group and 1-year group before maturity. I have also cut the sample by daily

quoting activity looking at the average of the top 10% of days and reporting the variances and covariances for the highest activity days for each contract in my sample.

I offer a more detailed look at the maturity effect, computing average bids and asks wavelet variances over 30 days, 60 days, 3 months, and 1 year before the contracts' maturity date considering only the days when the number of quotes is at least 5,000. To analyze the performance of high-price contract groups (such as the 76%–100% group), Table 2.4 reports the median variance of bids and asks, the median bid-ask wavelet covariance and the median wavelet correlations for the first quarter delivery contracts (EDH group) corresponding to time-scale levels 1 and level 9.[1] The Table illustrates that there is an increasing trend for the median variance of asks and bids and for the median wavelet covariance as the time to maturity decreases from 1 year to 30 days. For instance, for the March 2010 delivery (EDH0) Eurodollar future at timescale level 1, the median ask variance one year before maturity is 1.7864. A month before maturity it increases to 3.4744. For the median bid variance, it also increases as maturity approaches in the case of the March 2010 delivery contract; at timescale level 1 its evolution is given by 1.7574 (1 year to maturity), 2.03 (3 months to maturity), 2.32 (2 months to maturity), and 3.3829 (1 month to maturity). In economic terms, the excess variance from the bid side translates to around 6bp per quote (multiplying the median spread by the variance ratio). This trend is also observed in the case of the bid–ask covariance of this contact, which increases from 1.7495 to 3.4333 as time to maturity decreases from 1 year to 30 days. Moreover, the trend is observed in all the cases reported in the Table. Therefore, overall my empirical evidence supports Samuelson's maturity effect theory in that the bid and ask variances and covariance of the first quarter delivery contracts gradually increase

[1]The results for all nine level timescales are available in the online appendix.

Table 2.4: First Quarter Delivery Futures Contracts Median Results

RIC	Timescale	Asks variance				Bids variance				Bid–ask covariance				Correlations			
		30 days before	60 days before	3 months before	1 Year before	30 days before	60 days before	3 months before	1 Year before	30 days before	60 days before	3 months before	1 Year before	30 days before	60 days before	3 months before	1 Year before
EDH0	Level 1	3.4744	2.3066	2.0378	1.7864	3.3829	2.32	2.03	1.7574	3.4333	2.3054	1.9665	1.7495	0.6038	0.6688	0.6667	0.6711
		(10.1 s)	(4.5 s)	(2.6 s)	(1.4 s)	(8.9 s)	(4.3 s)	(2.5 s)	(1.3 s)	(9.5 s)	(4.4 s)	(2.6 s)	(1.3 s)	(9.5 s)	(4.4 s)	(2.6 s)	(1.3 s)
	Level 9	1.2728	1.2193	1.1406	1.0656	1.3946	1.3213	1.2367	1.0691	1.1029	1.0807	1.0492	1.0311	0.9495	0.9689	0.9735	0.9840
		(43.1 min)	(19.2 min)	(11.2 min)	(5.8 min)	(38.2 min)	(18.5 min)	(10.9 min)	(5.7 min)	(40.6 min)	(18.9 min)	(11.1 min)	(5.7 min)	(40.6 min)	(18.9 min)	(11.1 min)	(5.7 min)
EDH1	Level 1	3.7076	3.3471	2.0818	1.7418	3.7897	3.3539	2.0210	1.7527	3.7477	3.3491	2.0115	1.7361	0.6648	0.6616	0.6606	0.6679
		(11.6 s)	(5.4 s)	(3.2 s)	(1.9 s)	(12.6 s)	(5.4 s)	(3.2 s)	(1.9 s)	(12.1 s)	(5.4 s)	(3.2 s)	(1.9 s)	(12.1 s)	(5.4 s)	(3.2 s)	(1.9 s)
	Level 9	1.5263	1.4767	1.1700	1.0551	1.4762	1.6234	1.2859	1.0659	1.2883	1.1262	1.0448	1.0222	0.9699	0.9771	0.9783	0.9822
		(49.7 min)	(23.1 min)	(13.8 min)	(8.1 min)	(53.7 min)	(22.9 min)	(13.4 min)	(8.2 min)	(51.7 min)	(23 min)	(13.6 min)	(8.2 min)	(51.7 min)	(23 min)	(13.6 min)	(8.2 min)
EDH2	Level 1	2.4838	1.9626	1.9296	1.7581	2.5987	2.0013	1.9425	1.7515	2.2186	1.9482	1.9279	1.7481	0.6654	0.6626	0.6626	0.6710
		(14.3 s)	(6.5 s)	(4.2 s)	(2.9 s)	(13.1 s)	(6 s)	(3.9 s)	(2.7 s)	(13.7 s)	(6.3 s)	(4.1 s)	(2.8 s)	(13.7 s)	(6.3 s)	(4.1 s)	(2.8 s)
	Level 9	1.1007	1.0624	1.0624	1.0415	1.1471	1.0985	1.0922	1.0503	1.0461	1.0365	1.0339	1.0219	0.9789	0.9809	0.9812	0.9857
		(60.9 min)	(27.7 min)	(17.9 min)	(12.2 min)	(55.9 min)	(25.8 min)	(16.8 min)	(11.6 min)	(58.4 min)	(26.7 min)	(17.4 min)	(11.9 min)	(58.4 min)	(26.7 min)	(17.4 min)	(11.9 min)
EDH3	Level 1	2.5168	2.4636	1.9928	1.8368	2.2232	2.1235	2.0027	1.8500	2.2095	2.1183	1.9960	1.8520	0.6645	0.6632	0.6666	0.6660
		(12.4 s)	(6.3 s)	(4.5 s)	(3 s)	(11.6 s)	(6.4 s)	(4.4 s)	(2.9 s)	(12 s)	(6.3 s)	(4.4 s)	(3 s)	(12 s)	(6.3 s)	(4.4 s)	(3 s)
	Level 9	1.8433	1.1178	1.0996	1.0570	8.2292	1.1542	1.1235	1.0744	1.0860	1.0433	1.0339	1.0218	0.9833	0.9896	0.9899	0.9850
		(52.7 min)	(26.8 min)	(19.1 min)	(12.8 min)	(49.5 min)	(27.2 min)	(18.6 min)	(12.5 min)	(51.1 min)	(27 min)	(18.8 min)	(12.6 min)	(51.1 min)	(27 min)	(18.8 min)	(12.6 min)
EDH4	Level 1	2.5236	2.7287	2.5236	1.9458	2.5266	2.6237	2.5266	1.9420	2.5251	2.6687	2.0707	1.9382	0.6657	0.6625	0.6625	0.6562
		(9.5 s)	(7.3 s)	(4.8 s)	(3 s)	(11.8 s)	(7.4 s)	(5.1 s)	(3.2 s)	(10.6 s)	(7.4 s)	(5 s)	(3.1 s)	(10.6 s)	(7.4 s)	(5 s)	(3.1 s)
	Level 9	1.1074	1.1447	1.1074	1.0431	1.0906	1.1421	1.0906	1.0562	1.0268	1.0268	1.0265	1.0177	0.9913	0.9913	0.9911	0.9862
		(40.6 min)	(31.3 min)	(20.7 min)	(12.9 min)	(50.3 min)	(31.8 min)	(21.7 min)	(13.7 min)	(45.4 min)	(31.5 min)	(21.2 min)	(13.3 min)	(45.4 min)	(31.5 min)	(21.2 min)	(13.3 min)
EDH5	Level 1	2.2329	2.2088	2.1164	1.7707	2.2685	2.1735	2.1037	1.7746	2.2453	2.1733	2.1105	1.7685	0.6449	0.6448	0.6616	0.6820
		(16.1 s)	(9.9 s)	(5.9 s)	(5.7 s)	(16.1 s)	(9.9 s)	(5.9 s)	(5.7 s)	(16.1 s)	(9.9 s)	(5.9 s)	(5.7 s)	(16.1 s)	(9.9 s)	(5.9 s)	(5.7 s)
	Level 9	1.1708	1.1254	1.0960	1.0630	1.1539	1.1159	1.0895	1.0657	1.0905	1.0671	1.0593	1.0368	0.9698	0.9675	0.9760	0.9869
		(68.7 min)	(42.1 min)	(25.2 min)	(24.3 min)	(68.7 min)	(42.1 min)	(25.2 min)	(24.3 min)	(68.7 min)	(42.1 min)	(25.2 min)	(24.3 min)	(68.7 min)	(42.1 min)	(25.2 min)	(24.3 min)
EDH6	Level 1	2.1622	1.9414	1.8877	1.7362	2.1986	1.9565	1.8998	1.7338	2.1705	1.9404	1.8906	1.7293	0.6433	0.6639	0.6658	0.6807
		(21.2 s)	(9.4 s)	(5.7 s)	(3 s)	(21.2 s)	(9.4 s)	(5.7 s)	(3 s)	(21.2 s)	(9.4 s)	(5.7 s)	(3 s)	(21.2 s)	(9.4 s)	(5.7 s)	(3 s)
	Level 9	1.2575	1.1255	1.1089	1.0593	1.3131	1.1185	1.1132	1.0604	1.1399	1.0726	1.0622	1.0370	0.9524	0.9709	0.9748	0.9835
		(90.5 min)	(39.9 min)	(24.4 min)	(12.6 min)	(90.5 min)	(39.9 min)	(24.4 min)	(12.6 min)	(90.5 min)	(39.9 min)	(24.4 min)	(12.6 min)	(90.5 min)	(39.9 min)	(24.4 min)	(12.6 min)
EDH7	Level 1	2.0021	1.9941	1.9869	1.7165	2.0235	2.0181	2.0071	1.7125	2.0111	2.0051	2.0009	1.7080	0.6652	0.6621	0.6642	0.6729
		(17.9 s)	(8 s)	(5.4 s)	(2.7 s)	(17.9 s)	(8 s)	(5.4 s)	(2.7 s)	(17.9 s)	(8 s)	(5.4 s)	(2.7 s)	(17.9 s)	(8 s)	(5.4 s)	(2.7 s)
	Level 9	1.8697	1.3141	1.1952	1.0692	1.1353	1.1333	1.1373	1.0613	1.1047	1.0957	1.1019	1.0371	0.9812	0.9826	0.9826	0.9852
		(76.2 min)	(34.2 min)	(22.9 min)	(11.4 min)	(76.2 min)	(34.2 min)	(22.9 min)	(11.4 min)	(76.2 min)	(34.2 min)	(22.9 min)	(11.4 min)	(76.2 min)	(34.2 min)	(22.9 min)	(11.4 min)
EDH8	Level 1	1.8763	1.7369	1.7347	1.6506	1.8736	1.7315	1.7315	1.6513	1.8746	1.7271	1.7271	1.6493	0.6649	0.6724	0.6896	0.6754
		(9.1 s)	(4.3 s)	(3.1 s)	(1.7 s)	(9.1 s)	(4.3 s)	(3.1 s)	(1.7 s)	(9.1 s)	(4.3 s)	(3.1 s)	(1.7 s)	(9.1 s)	(4.3 s)	(3.1 s)	(1.7 s)
	Level 9	1.0591	1.0480	1.0452	1.0374	1.0592	1.0441	1.0441	1.0402	1.0283	1.0227	1.0212	1.0227	0.9816	0.9900	0.9900	0.9887
		(38.6 min)	(18.4 min)	(13.2 min)	(7.1 min)	(38.6 min)	(18.4 min)	(13.2 min)	(7.1 min)	(38.6 min)	(18.4 min)	(13.2 min)	(7.1 min)	(38.6 min)	(18.4 min)	(13.2 min)	(7.1 min)
EDH9	Level 1	1.7183	1.6637	1.6725	1.6393	1.7168	1.6721	1.6755	1.6406	1.7163	1.6656	1.6706	1.6381	0.6768	0.6768	0.6780	0.6796
		(3.5 s)	(1.9 s)	(1.4 s)	(1 s)	(3.5 s)	(1.9 s)	(1.4 s)	(1 s)	(3.5 s)	(1.9 s)	(1.4 s)	(1 s)	(3.5 s)	(1.9 s)	(1.4 s)	(1 s)
	Level 9	1.0354	1.0283	1.0314	1.0249	1.0357	1.0333	1.0337	1.0251	1.0178	1.0174	1.0175	1.0132	0.9864	0.9877	0.9890	0.9919
		(15.1 min)	(8.1 min)	(5.9 min)	(4.4 min)	(15.1 min)	(8.1 min)	(5.9 min)	(4.4 min)	(15.1 min)	(8.1 min)	(5.9 min)	(4.4 min)	(15.1 min)	(8.1 min)	(5.9 min)	(4.4 min)

Notes: This table reports EDH futures contracts median results for wavelet level 1 and level 9 (76–100% group) timescales for bid variance, ask variance, bid-ask wavelet covariance and wavelet correlations for high-price-level contracts for EDH0 to EDH9. The study eliminates days with less than 5,000 observations, and compares different periods before the settlement date: 30 days, 60 days, 3 months and 1 year before Eurodollar contract maturity date. Contracts are also divided into four different price groups: 0–25% prices, 26–50% prices, 51–75% prices and 76–100% prices. For these sub-groups, I calculate all the timescales for nine levels of EDH contracts.

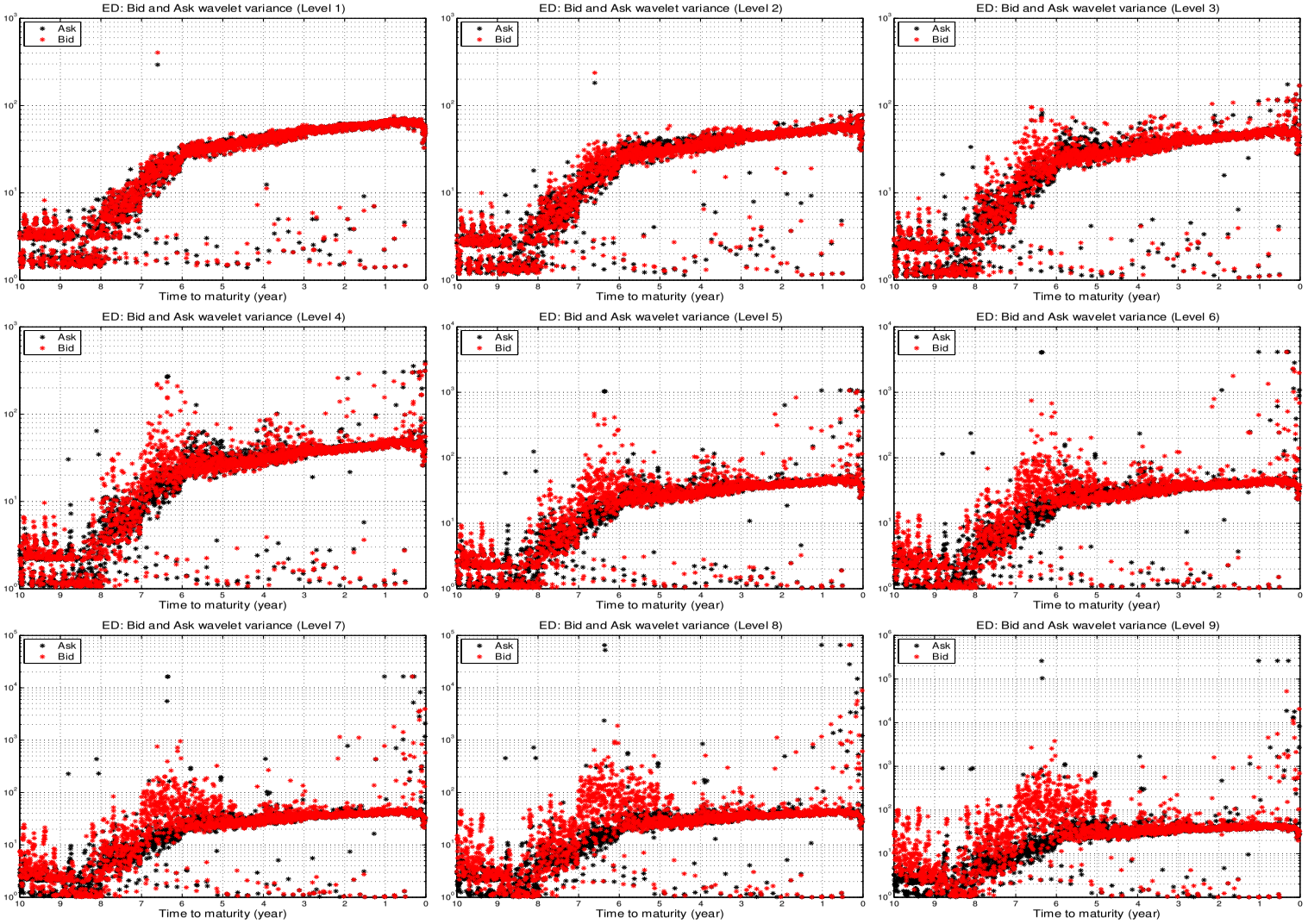


Figure 2.5: The Term Structure of 40 Eurodollar Future Contracts Wavelet Bid and Ask Variance for Timescale Levels 1-9

Notes: The black markers reports the ask variance ratios and red markets report the bid variance ratios. I eliminate days with less than 5,000 bid or ask quotes. Level 1 refers to the highest frequency timestamp and level 9 is the lowest frequency timestamp.

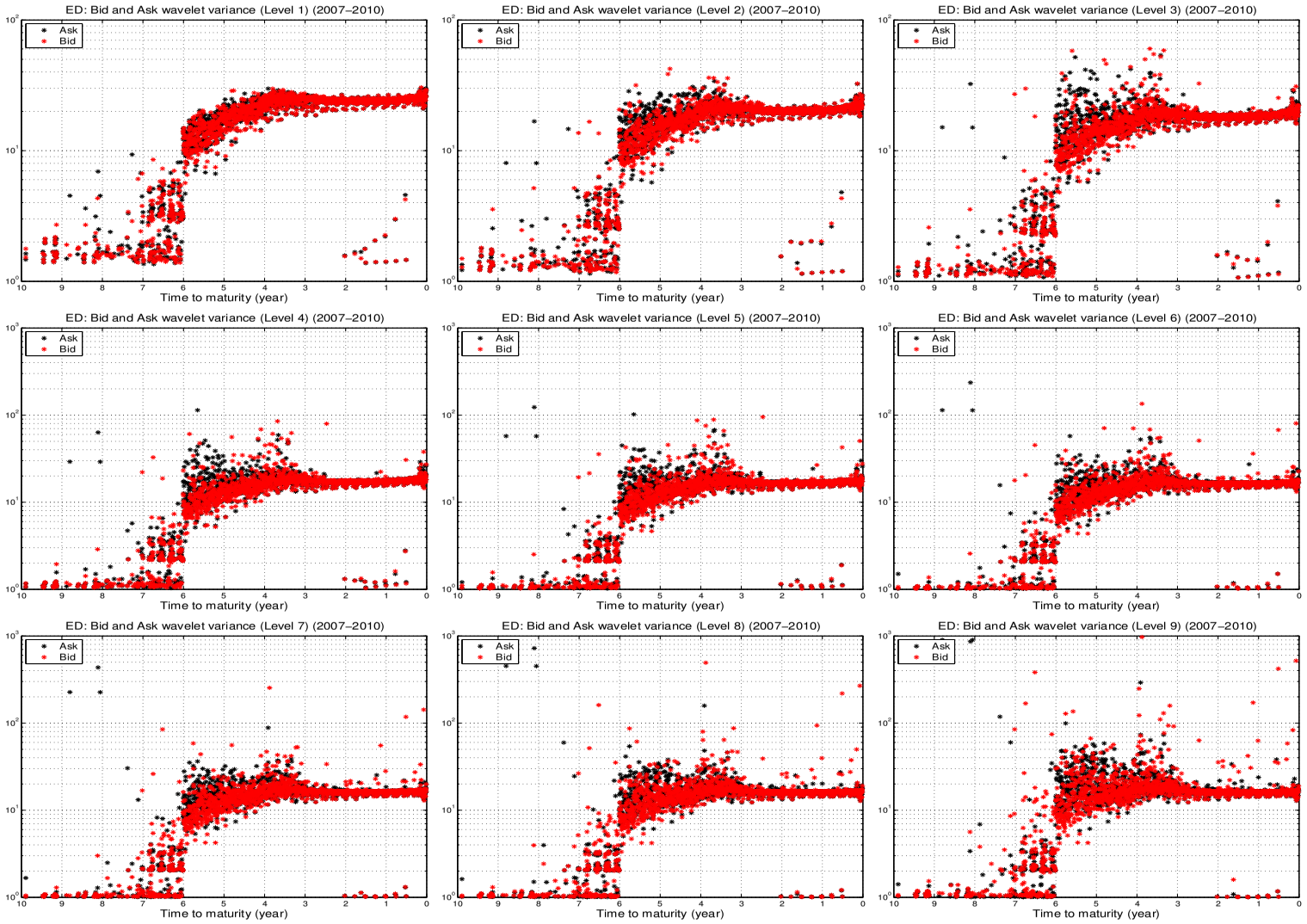


Figure 2.6: The Term Structure of 40 Eurodollar Future Contracts Wavelet Bid and Ask Variance for Timescale Levels 1-9 (2007 - 2010)

Notes: The black markers reports the ask variance ratios and red markets report the bid variance ratios. I eliminate days with less than 5,000 bid or ask quotes. Level 1 refers to the highest frequency timestamp and level 9 is the lowest frequency timestamp.

as maturity approaches.

Bessembinder et al. [1996] point out that the maturity effect is not observed in all markets, especially for financial assets settling to cash as opposed commodities with a physical delivery, such as crude oil or agricultural commodities. Kalev and Duong [2008] do not find support for the Samuelson hypothesis in CME Eurodollar future markets. However, my empirical evidence supports Samuelson's maturity effect theory. From the 1-year window to 30 days before maturity, the EDH bid and ask variances and covariance gradually increase. This indicates that the Eurodollar Futures price volatilities increase when the Eurodollar contracts are close to expiration. At this moment it is important to highlight that Spoof quoting is common in the equity markets and is used to disguise positions within the order book [Hasbrouck, 2014]. Moreover, the recent regulatory interventions in the reference rate indicate that there is quite likely to be substantial information asymmetry across market participants in the Eurodollar futures market. The complete short-term structure of interest rates is updated during a daily LIBOR fixing, this will in turn update the implied impact on the forward curve approximated by the variously maturing Eurodollar futures. As such, to manipulate the price of a contract one needs to manipulate the appropriate point in the yield curve rather than just the 3-month LIBOR rate. Full identification and comparison of trading activity during periods of heightened LIBOR manipulation is therefore very difficult. The Eurodollar rate is intimately linked to the fixed to floating swap market and reference rate manipulation can benefit a counter-party in many different legs of a given swap-Eurodollar deal. So there are many steps that may provide systematic arbitrage opportunities that cannot be identified by this type of analysis alone. This is outside the scope of this study.

Variance ratios for EDH4 and EDH9 median value at level 1 and level 9 are quite

high, in fact, they rise to levels similar to that found in the equity markets. For the high price level group (76–100%), the EDH4 median value in Table 2.4 indicates a small jump from the 30-day window to the 60-day window. The median of asks variance in the 30-day window is 2.5236 at level 1, while that for 60 days is 2.7287. EDH4 bid variance and wavelet covariance also increase by 0.0971 and 0.1436, respectively. In comparison, small jumps of EDH9 occurred between the 60-day window and the 3-month window at level 1. Moreover, the median of EDH9 wavelet covariance at level 1 only increases from 1.6656 to 1.6706 between the 60-day window and the 3-month window.

The median of bid-ask wavelet correlations, however, shows a generally increasing trend as the contracts approach maturity. Hasbrouck [2014] argues that the better functioning markets should have nearly perfectly correlated bids and asks and this is supported by several recent theoretical contributions, see Hansen and Lunde [2006] as an example. However, my results are in strong accordance without found Hasbrouck [2014]; which indicates that at a timescale of 800 ms or lower, the bid/ask variance correlation is below 0.6 for 80% of the US equities in 2011; only those equities in the highest trading volume quintiles have 0.72. The wavelet correlations of ED contracts in Table 2.4 are relatively higher than Hasbrouck's results of the US equities. Table 2.4 shows the median value of correlation increases from 0.6038 (30 days window) to 0.6711 (one year window) at timescale level 1. Actually, the correlations of all EDH contracts at level 1 are larger than 0.6 and at level 9 are larger than 0.9 and close to 1. If the bid and ask variation were perfectly correlated, public information revelation would shift both bids and asks prices by the same amount.

In Figures 2.5 and 2.6, I present the term structure of the wavelet bid and ask variance for the 40 Eurodollar future contracts. To construct the Figures, I first

perform a nine time-scale level wavelet multi-resolution analysis considering only the days when the number of quotes is at least 5,000.[1] Then, I compute the sum of bids and asks wavelet variances for all the contracts.

To provide a complementary understanding on the maturity effects in Figures 2.5 - 2.6, I choose to plot the sum of wavelet variances based on the time to maturity and not on the actual transaction dates used by Hendershott et al. [2011] and Hasbrouck [2014]. Given that each Eurodollar futures contract has a tenor of 10-year, my procedure implies plotting the term structure from 10 to 0 years to maturity. Figure 2.5 exhibits the results using all the years in my sample (1996 – 2013), while Figure 2.6 is focused on the period of the recent financial crisis (2007 – 2010). Wavelet variance plots for March, June, September and December delivery contracts and bid-ask variance plots of each single contract can be found in the online appendix.

Both Figures show that during the first two years the quoting variance rate is relatively stable at all the timescale levels, indicating that the price variability for every contract is less vigorous at the beginning of its lifecycle. From eight years before maturity, the variance ratios start to increase reaching their maximum value around two years before maturity. This change is consistent with the maturity effect behavior, characterized by monotone increases in price variability as the contract approaches maturity. Interestingly, around one week before maturity the quoting variance suddenly drops. Figure 2.5 also shows the asymmetry between the bids and asks price variability on both the degree of changes and the movement directions in all the nine levels timescales.

[1] This is a relatively arbitrary choice. The main reason for it is that 5,000 is the minimum number of updates for the variation in quotes needed to perform a nine levels wavelet decomposition. Eurodollar Futures quotes tend to be less frequent than an actively traded stock, albeit with a far larger volume attached to each individual quote. 5,000 updates result in no singularities within my dataset.

The maturity effect can be clearly seen in Figure 2.6, which is focused on the period of the recent financial crisis from 2007 to 2010. This plot illustrates ED price volatility during the financial crisis. Consistent with the maturity effect conjecture, quoting volatility maintains a low level at the beginning of time to maturity, and then monotonically increases since six years before maturity. Compared with the variance of full sample (see Figure 2.5), there is less vigorous during the first four years. Around the last two years, it shows the similar pattern with Figure 2.5. Both bids and asks variances reach the highest level, but substantially decrease at 1 week approaching maturity.

To record the time series variation, I calculate year-long averages of the daily bid-ask wavelets variance, covariance and correlations for all ED contracts. Tables 2.5 and 2.6 report the yearly median value of EDH0 and EDH8. Yearly results for other March, June, September and December delivery contracts are available in the online appendix and provide materially similar results. Table 2.5 presents the median bids and asks variance, the covariance and the wavelet correlations of the EDH0 contract at all timescales by year from 2006 to 2013.[1] There are no results in 2011 given that EDH0 was delivered in March 2010 and then rolled over to a new 10-year period contract, so that there are virtually no trades or quotes. As observed in Table 2.5, the ask and the bid variances of EDH0 show obvious increases from 2009 to 2010, therefore also providing evidence in favor of the maturity effect theory that the variability of contract prices should increase monotonically toward the maturity date [Samuelson, 1965]. However, the median results of both the bid and the ask variance decrease between 2007 and 2008 which may be a consequence of the recent subprime mortgage market collapse, which led investors to be more careful about what they invested in. This interpretation is further supported by the dramatic rise experienced by the ask and the bid

[1] Years before 2005 have been eliminated due to the lack of observations.

variances during the 2009-2010 period, which seems to be the result of a large market correction.

Table 2.6 displays the bid and ask variance, the covariance and the wavelet correlations at timescales level 1 to level 9 from 2005 to 2013. It reports the same information for the EDH8 contract, a potentially informative contract, due to its maturity date on March 2008, lies precisely over the dates of the March 2008 Bear Stearns bankruptcy. Interestingly, although EDH8 was delivered in March 2008 and then began a new 10-year period of contract there are no year gaps even after its maturity date in 2009. The results reported in Table 2.6 show that in the case of EDH8 the variances increased slightly moving to 2008 although without the sudden large jumps expected in the maturity effect theory of Samuelson [1965]. I believe this might be an effect of the financial crisis and of investors lack of confidence in the Eurodollar Futures market as, coinciding with the beginning of the second EDH8 contract period, the variances increased from 2009 to 2010.

Yet another interesting feature my sample allow me to do is to construct a set of annual Eurodollar Future forward rates mapping to the forward curve. To construct the set, I group ED futures by delivery date and, considering only the days when the number of quotes is at least 5,000, I compute bids and asks wavelet variances for each ED futures group over 30 days, 60 days, 3 months, and 1 year before the contracts' maturity date. Then, I separate contracts according to their prices and average the variances across price quartiles using each of the four time horizons listed above. The results, reported in Tables 2.7 and 2.8, allow me to compare the Eurodollar futures market with the US equity market in terms of fundamental and transient volatility and therefore in terms of decoupling of bid and ask sides of the book over time.

A number of elements in Tables 2.7 and 2.8 are worth mentioning. First, the

Table 2.5: EDH0 Median Annual Variance Ratios with Time to Maturity

EDH0	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**EDH0**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
				Ask variance									**Bid variance**				
Level 1	1.6028	1.5800	1.5186	1.5487	1.8282	-	1.6690	1.5920	Level 1	1.6606	1.5830	1.5178	1.5451	1.7656	-	1.6655	1.5764
Level 2	1.3694	1.3423	1.2662	1.3056	1.6067	-	1.3970	1.3308	Level 2	1.4049	1.3348	1.2671	1.3059	1.5711	-	1.4502	1.3457
Level 3	1.2178	1.1835	1.1362	1.1830	1.4267	-	1.2763	1.1993	Level 3	1.2158	1.1866	1.1362	1.1864	1.4124	-	1.2865	1.2233
Level 4	1.1196	1.1051	1.0710	1.1197	1.2872	-	1.2006	1.1300	Level 4	1.1266	1.1058	1.0716	1.1211	1.2752	-	1.2749	1.1571
Level 5	1.0726	1.0650	1.0401	1.0838	1.2046	-	1.1829	1.1023	Level 5	1.0848	1.0652	1.0405	1.0838	1.1837	-	1.3041	1.1310
Level 6	1.0516	1.0435	1.0250	1.0576	1.1425	-	1.1670	1.0839	Level 6	1.0593	1.0442	1.0253	1.0567	1.1288	-	1.2901	1.1236
Level 7	1.0374	1.0318	1.0171	1.0391	1.1050	-	1.1246	1.0738	Level 7	1.0465	1.0320	1.0179	1.0386	1.1001	-	1.2373	1.1264
Level 8	1.0307	1.0250	1.0131	1.0272	1.0719	-	1.0903	1.0624	Level 8	1.0399	1.0256	1.0134	1.0276	1.0719	-	1.1888	1.1328
Level 9	1.0338	1.0234	1.0121	1.0241	1.0621	-	1.0732	1.0618	Level 9	1.0408	1.0255	1.0121	1.0236	1.0718	-	1.2142	1.2205
				Bid–ask covariance									**Correlations**				
Level 1	1.6005	1.5623	1.5158	1.5435	1.7851	-	1.5726	1.5217	Level 1	0.6022	0.6317	0.6589	0.6472	0.5664	-	0.6004	0.6344
Level 2	1.3308	1.3206	1.2603	1.2994	1.5564	-	1.3093	1.2481	Level 2	0.7118	0.7492	0.7892	0.7658	0.6365	-	0.6895	0.7431
Level 3	1.1607	1.1624	1.1301	1.1728	1.3512	-	1.1431	1.1127	Level 3	0.8225	0.8428	0.8802	0.8429	0.7080	-	0.7773	0.8175
Level 4	1.0789	1.0775	1.0636	1.1091	1.2285	-	1.0794	1.0464	Level 4	0.8877	0.9043	0.9332	0.8920	0.7842	-	0.7844	0.8642
Level 5	1.0384	1.0371	1.0313	1.0704	1.1563	-	1.0623	1.0161	Level 5	0.9218	0.9388	0.9611	0.9227	0.8448	-	0.7668	0.8842
Level 6	1.0172	1.0184	1.0163	1.0437	1.1032	-	1.0533	1.0000	Level 6	0.9440	0.9577	0.9754	0.9464	0.8859	-	0.7752	0.8900
Level 7	1.0080	1.0092	1.0087	1.0270	1.0645	-	1.0334	0.9895	Level 7	0.9556	0.9689	0.9824	0.9629	0.9090	-	0.8082	0.8878
Level 8	1.0033	1.0047	1.0047	1.0172	1.0402	-	1.0124	0.9854	Level 8	0.9616	0.9751	0.9868	0.9731	0.9330	-	0.8412	0.8828
Level 9	1.0013	1.0022	1.0031	1.0125	1.0274	-	1.0071	0.9878	Level 9	0.9608	0.9751	0.9881	0.9769	0.9330	-	0.8236	0.8193

Notes: This table shows the yearly median bid and ask variance, covariance and wavelet correlations at all timescales for EDH0 futures contracts from 2006 to 2013. The years prior to 2006 are blank due to the lack of trading activity. As EDH0 was delivered in March 2010 and rolled over to a new 10-year contract, results in 2011 are also not available due to being at the beginning of a new period EDH0 contract.

Table 2.6: EDH8 Median Annual Variance Ratios with Time to Maturity

EDH8	2005	2006	2007	2008	2009	2010	2011	2012	2013
Ask variance									
Level 1	1.5738	1.5377	1.5423	1.5547	1.6127	1.8878	1.7916	1.5277	1.5162
Level 2	1.3181	1.2890	1.2926	1.2784	1.3280	1.8168	1.5561	1.2894	1.2734
Level 3	1.1901	1.1698	1.1665	1.1512	1.1809	1.6217	1.3804	1.1671	1.1551
Level 4	1.1180	1.1137	1.1035	1.0846	1.1075	1.3149	1.3071	1.1084	1.0948
Level 5	1.0788	1.0818	1.0706	1.0528	1.0603	1.1646	1.2626	1.0815	1.0619
Level 6	1.0559	1.0569	1.0505	1.0347	1.0397	1.0830	1.2545	1.0606	1.0402
Level 7	1.0438	1.0392	1.0352	1.0242	1.0366	1.0442	1.2226	1.0460	1.0254
Level 8	1.0362	1.0288	1.0261	1.0185	1.0462	1.0282	1.2279	1.0355	1.0168
Level 9	1.0367	1.0262	1.0221	1.0193	1.0791	1.0243	1.1901	1.0322	1.0132

EDH8	2005	2006	2007	2008	2009	2010	2011	2012	2013
Bid variance									
Level 1	1.5770	1.5379	1.5424	1.5517	1.5867	1.7845	1.6795	1.5647	1.5158
Level 2	1.3259	1.2898	1.2919	1.2776	1.3313	1.4601	1.4096	1.3304	1.2657
Level 3	1.1902	1.1675	1.1677	1.1469	1.1672	1.2458	1.2410	1.1959	1.1467
Level 4	1.1218	1.1129	1.1055	1.0844	1.0924	1.1414	1.1498	1.1307	1.0891
Level 5	1.0908	1.0778	1.0712	1.0506	1.0529	1.1181	1.0943	1.1109	1.0579
Level 6	1.0680	1.0541	1.0510	1.0319	1.0368	1.1220	1.0685	1.1075	1.0381
Level 7	1.0530	1.0374	1.0362	1.0230	1.0375	1.1673	1.0557	1.0902	1.0256
Level 8	1.0495	1.0277	1.0262	1.0184	1.0487	1.2708	1.0488	1.1050	1.0191
Level 9	1.0483	1.0241	1.0227	1.0193	1.0884	1.5163	1.0734	1.1354	1.0180

EDH8	2005	2006	2007	2008	2009	2010	2011	2012	2013
Bid–ask covariance									
Level 1	1.5582	1.5308	1.5407	1.5511	1.5496	1.7775	1.5759	1.4935	1.5034
Level 2	1.2944	1.2784	1.2893	1.2762	1.2346	1.5000	1.2959	1.2401	1.2539
Level 3	1.1452	1.1544	1.1611	1.1459	1.1137	1.3088	1.1723	1.1083	1.1247
Level 4	1.0680	1.0920	1.0979	1.0813	1.0397	1.1205	1.0756	1.0509	1.0640
Level 5	1.0310	1.0581	1.0635	1.0465	1.0219	1.0237	1.0401	1.0219	1.0320
Level 6	1.0142	1.0344	1.0430	1.0279	1.0131	1.0108	1.0199	1.0070	1.0159
Level 7	1.0067	1.0201	1.0278	1.0179	1.0078	1.0045	1.0110	0.9998	1.0077
Level 8	1.0036	1.0127	1.0182	1.0121	1.0053	0.9997	1.0051	0.9972	1.0035
Level 9	1.0017	1.0085	1.0134	1.0099	1.0024	0.9989	1.0024	0.9974	1.0018

EDH8	2005	2006	2007	2008	2009	2010	2011	2012	2013
Correlations									
Level 1	0.6341	0.6503	0.6483	0.6445	0.6303	0.5604	0.5954	0.6391	0.6597
Level 2	0.7542	0.7753	0.7741	0.7827	0.7512	0.6849	0.7094	0.7517	0.7900
Level 3	0.8402	0.8566	0.8564	0.8719	0.8569	0.8027	0.8058	0.8362	0.8720
Level 4	0.8914	0.8986	0.9046	0.9222	0.9154	0.8761	0.8697	0.8844	0.9182
Level 5	0.9168	0.9278	0.9335	0.9518	0.9498	0.8944	0.9138	0.9002	0.9452
Level 6	0.9363	0.9487	0.9515	0.9691	0.9645	0.8913	0.9359	0.9029	0.9633
Level 7	0.9497	0.9639	0.9651	0.9775	0.9638	0.8567	0.9472	0.9172	0.9750
Level 8	0.9529	0.9731	0.9744	0.9820	0.9535	0.7869	0.9535	0.9050	0.9813
Level 9	0.9539	0.9765	0.9778	0.9810	0.9189	0.6595	0.9316	0.8808	0.9823

Notes: This table describes that the yearly median bid and ask variance, covariance and wavelet correlations at all timescales for EDH8 futures contracts from 2005 to 2013. Years before 2005 are eliminated because of the lack of trading activity. As EDH8 was delivered in March 2008, it rolls over to a new 10-year period contract from 2008 to 2018.

Tables show that the bid variances for all pricing groups and at all time-scales are relatively similar to the ask variances meaning that, unlike Hasbrouck's findings for the US equity markets, the effects of both best bid and best ask appear to be about the same for Eurodollar Futures. My results also suggest that the higher price groups tend to have the higher variance and wavelet bid-ask correlations. For example, at time-scale level 1 the average of second quarter delivery contracts (EDM) asks variance in the 30 days before maturity rises from 1.5223 in the lowest price group to 2.6613 in the highest price group at the 14.3 seconds timescale. In the same time-scale level, the average of second quarter delivery contracts bids variance in the 30-day period also increases from 1.5157 to 2.638 when moving to higher price groups. Note also that both bid and ask variances rise with shorter timescales. These results highlight the differences between the Eurodollar Futures market and the US equity market where Hasbrouck [2014] finds that the volatility of the lowest dollar volume quintile is much higher than that of the highest quintile. Table 2.7 also shows that the volatilities of both the best-bid and the best-ask increase before approaching maturity. For the 76-100% price group at timescale level 1, the ask variance of fourth quarter delivery contracts (EDZ) 1 year before maturity is 1.7641, while its ask variance 30 days before maturity increases to 2.6613 at timescale level 1 for the highest price level group.

The average results of bid–ask covariance and correlations for EDM, EDU and EDZ are recorded in Table 2.8. The volatilities of bids and asks of every maturity-type contract maintain positive covariances and correlations at all timescales. For instance, the average wavelet correlations are larger than 0.8 for all four price levels at timescale level 9. This suggests that the bid volatility and ask volatility are in significant correlation just prior to the ED Future maturity date.

To explore the transient nature of Eurodollar future contracts quoting frequency,

Table 2.7: The Average of Median Bid and Ask Variances of Second, Third and Fourth Quarter Delivery Contracts

		EDM					EDU					EDZ				
		Time scale	**0%– 25%**	**26%– 50%**	**51%– 75%**	**76%– 100%**	**Time scale**	**0%– 25%**	**26%– 50%**	**51%– 75%**	**76%– 100%**	**Time scale**	**0%– 25%**	**26%– 50%**	**51%– 75%**	**76%– 100%**
Ask variance																
	Level 1	14.3 s	1.5223	1.6642	1.8574	2.6613	13.2 s	1.4990	1.6217	1.8425	2.4568	12.6 s	1.5028	1.6034	1.7848	2.4099
30	Level 3	57 s	1.1933	1.3057	1.4947	1.9823	52.6 s	1.1685	1.2796	1.5151	2.0938	50.2 s	1.1675	1.2607	1.4143	2.1764
days	Level 5	3.8 min	1.0755	1.1408	1.2482	1.4973	3.5 min	1.0597	1.1122	1.2255	1.6027	3.3 min	1.0628	1.1286	1.1783	1.6865
before	Level 7	15.2 min	1.0341	1.0664	1.1286	1.3061	14 min	1.0268	1.0546	1.0971	1.2682	13.4 min	1.0273	1.0526	1.0833	1.3598
	Level 9	60.8 min	1.0248	1.0447	1.0951	1.2243	56.2 min	1.0169	1.0383	1.0650	1.2001	53.6 min	1.0156	1.0344	1.0606	1.2638
	Level 1	6.5 s	1.5054	1.6275	1.7590	2.2335	6.1 s	1.5057	1.6100	1.7338	2.1030	6.3 s	1.5127	1.6108	1.7462	2.2569
60	Level 3	26.2 s	1.1593	1.2737	1.4070	2.2219	24.5 s	1.1737	1.2618	1.3729	1.8134	25.3 s	1.1598	1.2598	1.3674	1.8980
days	Level 5	1.7 min	1.0577	1.1160	1.1921	1.7756	1.6 min	1.0645	1.1167	1.1713	1.3637	1.7 min	1.0606	1.1187	1.1773	1.5473
before	Level 7	7 min	1.0244	1.0466	1.0871	1.2994	6.5 min	1.0270	1.0485	1.0768	1.1874	6.8 min	1.0254	1.0508	1.0797	1.2554
	Level 9	27.9 min	1.0155	1.0296	1.0546	1.1944	26.1 min	1.0164	1.0302	1.0460	1.1161	27 min	1.0142	1.0297	1.0493	1.1784
	Level 1	4 s	1.5102	1.6067	1.7126	2.0749	3.7 s	1.5066	1.6031	1.7035	1.9746	3.7 s	1.4994	1.5889	1.6881	1.9669
3	Level 3	15.9 s	1.1660	1.2450	1.3593	1.6802	14.8 s	1.1764	1.2508	1.3356	1.5910	14.9 s	1.1600	1.2357	1.3286	1.5961
months	Level 5	1.1 min	1.0634	1.1054	1.1662	1.3264	59.3 s	1.0651	1.1090	1.1585	1.3047	59.6 s	1.0613	1.1024	1.1542	1.3035
before	Level 7	4.2 min	1.0244	1.0457	1.0735	1.1632	4 min	1.0274	1.0467	1.0716	1.1386	4 min	1.0248	1.0444	1.0695	1.1408
	Level 9	17 min	1.0149	1.0260	1.0437	1.1113	15.8 min	1.0161	1.0277	1.0419	1.0901	15.9 min	1.0132	1.0254	1.0410	1.0937
	Level 1	2.5 s	1.4876	1.5582	1.6287	1.7641	2.3 s	1.4907	1.5598	1.6274	1.7653	2.2 s	1.4907	1.5524	1.6146	1.7516
1	Level 3	9.9 s	1.1493	1.2022	1.2592	1.3925	9.3 s	1.1478	1.2042	1.2613	1.3890	8.9 s	1.1474	1.2001	1.2552	1.3741
Year	Level 5	39.8 s	1.0560	1.0871	1.1189	1.1909	37.1 s	1.0551	1.0865	1.1182	1.1900	35.8 s	1.0556	1.0851	1.1169	1.1863
before	Level 7	2.7 min	1.0227	1.0379	1.0528	1.0876	2.5 min	1.0225	1.0377	1.0526	1.0881	2.4 min	1.0231	1.0370	1.0520	1.0840
	Level 9	10.6 min	1.0130	1.0213	1.0303	1.0527	9.9 min	1.0126	1.0213	1.0301	1.0511	9.5 min	1.0126	1.0209	1.0295	1.0482
Bid variance																
	Level 1	14.1 s	1.5157	1.6654	1.8705	2.6380	13 s	1.4997	1.6464	1.8345	2.4982	12.4 s	1.5000	1.6019	1.7712	2.3790
30	Level 3	56.3 s	1.1913	1.2876	1.4552	1.8969	51.9 s	1.1695	1.2755	1.5179	2.1056	49.6 s	1.1674	1.2546	1.4070	2.4299
days	Level 5	3.8 min	1.0723	1.1372	1.2206	1.4300	3.5 min	1.0604	1.1048	1.2223	1.6326	3.3 min	1.0602	1.1220	1.1678	1.9648
before	Level 7	15 min	1.0324	1.0621	1.1311	1.2730	13.8 min	1.0269	1.0527	1.1045	1.3278	13.2 min	1.0259	1.0527	1.0848	1.5653
	Level 9	60.1 min	1.0251	1.0548	1.1111	1.2462	55.4 min	1.0192	1.0417	1.0816	1.2733	52.9 min	1.0179	1.0359	1.0642	1.4017
	Level 1	6.5 s	1.4972	1.6197	1.7574	2.2457	6 s	1.5050	1.6061	1.7353	2.0939	6 s	1.5102	1.6057	1.7462	2.2715
60	Level 3	26 s	1.1522	1.2675	1.3930	2.1705	24 s	1.1693	1.2524	1.3638	1.8242	24 s	1.1573	1.2539	1.3653	2.0684
days	Level 5	1.7 min	1.0559	1.1075	1.1896	1.7243	1.6 min	1.0620	1.1098	1.1699	1.3784	1.6 min	1.0590	1.1112	1.1641	1.7009
before	Level 7	6.9 min	1.0237	1.0485	1.0885	1.2854	6.4 min	1.0270	1.0489	1.0774	1.1999	6.4 min	1.0248	1.0509	1.0801	1.3820
	Level 9	27.7 min	1.0158	1.0331	1.0636	1.2085	25.6 min	1.0180	1.0340	1.0546	1.1502	25.6 min	1.0160	1.0327	1.0537	1.2366
	Level 1	4 s	1.5008	1.6018	1.7152	2.0695	3.6 s	1.5074	1.6038	1.7061	1.9712	3.6 s	1.5018	1.5883	1.6826	1.9272
3	Level 3	16 s	1.1617	1.2437	1.3392	1.6584	14.5 s	1.1731	1.2453	1.3335	1.6014	14.4 s	1.1572	1.2346	1.3213	1.5712
months	Level 5	1.1 min	1.0591	1.1020	1.1572	1.3139	58.1 s	1.0635	1.1035	1.1535	1.2971	57.4 s	1.0586	1.0972	1.1481	1.2812
before	Level 7	4.3 min	1.0248	1.0452	1.0719	1.1671	3.9 min	1.0269	1.0463	1.0710	1.1450	3.8 min	1.0250	1.0442	1.0676	1.1427
	Level 9	17 min	1.0159	1.0286	1.0475	1.1285	15.5 min	1.0177	1.0300	1.0465	1.1031	15.3 min	1.0147	1.0274	1.0434	1.0956
	Level 1	2.5 s	1.4867	1.5557	1.6232	1.7674	2.3 s	1.4867	1.5572	1.6231	1.7631	2.2 s	1.4894	1.5518	1.6122	1.7495
1	Level 3	10 s	1.1467	1.1989	1.2523	1.3812	9.2 s	1.1456	1.1977	1.2544	1.3771	8.8 s	1.1445	1.1960	1.2487	1.3609
Year	Level 5	39.9 s	1.0547	1.0842	1.1137	1.1797	36.7 s	1.0534	1.0831	1.1121	1.1808	35.1 s	1.0537	1.0814	1.1111	1.1772
before	Level 7	2.7 min	1.0224	1.0373	1.0517	1.0863	2.4 min	1.0221	1.0372	1.0523	1.0876	2.3 min	1.0228	1.0366	1.0514	1.0836
	Level 9	10.6 min	1.0138	1.0230	1.0323	1.0581	9.8 min	1.0138	1.0229	1.0326	1.0601	9.4 min	1.0137	1.0226	1.0324	1.0554

Notes: This table illustrates the average of EDM, EDU and EDZ median ask and bid variances within different timezones 30 days, 60 days, 3 months and 1 year before Eurodollar contract maturity date. Each group contains 10 Eurodollar contracts with the same delivery month in different years: EDM0, EDM1, ..., EDM9, with average results for all four price levels (0–25% price group, 26–50% price group, 51–75% price group and 76–100% price group).

Table 2.8: The Average Performance of Median Bid–ask Covariance and Correlations of Second, Third and Fourth Quarter Delivery Contracts

		EDM					EDU					EDZ				
		Time scale	**0%–25%**	**26%–50%**	**51%–75%**	**76%–100%**	**Time scale**	**0%–25%**	**26%–50%**	**51%–75%**	**76%–100%**	**Time scale**	**0%–25%**	**26%–50%**	**51%–75%**	**76%–100%**
Bid–ask covariance																
	Level 1	14.2 s	1.5191	1.6601	1.8598	2.6410	13.1 s	1.4923	1.6297	1.8328	2.4593	12.5 s	1.4972	1.5997	1.7654	2.3682
30	Level 3	56.7 s	1.1899	1.2870	1.4519	1.8888	52.3 s	1.1534	1.2605	1.4899	2.0638	49.9 s	1.1611	1.2474	1.4017	2.2038
days	Level 5	3.8 min	1.0676	1.1215	1.1930	1.3660	3.5 min	1.0528	1.0918	1.2017	1.5449	3.3 min	1.0553	1.1088	1.1514	1.6957
before	Level 7	15.1 min	1.0254	1.0443	1.0813	1.1948	13.9 min	1.0189	1.0362	1.0711	1.2060	13.3 min	1.0206	1.0379	1.0599	1.3332
	Level 9	60.4 min	1.0128	1.0251	1.0489	1.1089	55.8 min	1.0095	1.0179	1.0339	1.1121	53.2 min	1.0090	1.0175	1.0318	1.1547
	Level 1	6.5 s	1.4976	1.6178	1.7494	2.1800	6.1 s	1.5005	1.6023	1.7287	2.0885	6.2 s	1.5094	1.6044	1.7409	2.2592
60	Level 3	26.1 s	1.1515	1.2612	1.3716	1.7853	24.2 s	1.1640	1.2462	1.3432	1.7970	24.7 s	1.1527	1.2496	1.3496	1.9097
days	Level 5	1.7 min	1.0521	1.0983	1.1555	1.3276	1.6 min	1.0560	1.0959	1.1463	1.3223	1.6 min	1.0538	1.1043	1.1446	1.5317
before	Level 7	7 min	1.0175	1.0338	1.0600	1.1283	6.5 min	1.0194	1.0355	1.0559	1.1390	6.6 min	1.0197	1.0389	1.0591	1.2172
	Level 9	27.8 min	1.0078	1.0151	1.0250	1.0695	25.9 min	1.0084	1.0162	1.0250	1.0611	26.3 min	1.0083	1.0168	1.0271	1.0863
	Level 1	4 s	1.5013	1.5981	1.7043	2.0467	3.7 s	1.5034	1.5974	1.6999	1.9644	3.7 s	1.4987	1.5857	1.6813	1.9495
3	Level 3	15.9 s	1.1579	1.2349	1.3245	1.6043	14.7 s	1.1663	1.2384	1.3194	1.5663	14.6 s	1.1533	1.2282	1.3124	1.5653
months	Level 5	1.1 min	1.0558	1.0919	1.1389	1.2673	58.7 s	1.0581	1.0921	1.1332	1.2600	58.5 s	1.0541	1.0905	1.1344	1.2527
before	Level 7	4.2 min	1.0186	1.0329	1.0512	1.1088	3.9 min	1.0201	1.0346	1.0519	1.1005	3.9 min	1.0199	1.0344	1.0513	1.0997
	Level 9	17 min	1.0082	1.0142	1.0222	1.0538	15.7 min	1.0083	1.0149	1.0219	1.0462	15.6 min	1.0081	1.0151	1.0225	1.0448
	Level 1	2.5 s	1.4854	1.5521	1.6203	1.7561	2.3 s	1.4856	1.5543	1.6188	1.7559	2.2 s	1.4873	1.5466	1.6071	1.7427
1	Level 3	10 s	1.1423	1.1911	1.2432	1.3586	9.2 s	1.1410	1.1915	1.2460	1.3602	8.9 s	1.1393	1.1881	1.2388	1.3456
Year	Level 5	39.8 s	1.0493	1.0746	1.1010	1.1562	36.9 s	1.0478	1.0745	1.1002	1.1574	35.4 s	1.0483	1.0726	1.0991	1.1556
before	Level 7	2.7 min	1.0171	1.0282	1.0383	1.0616	2.5 min	1.0169	1.0281	1.0387	1.0626	2.4 min	1.0171	1.0276	1.0385	1.0606
	Level 9	10.6 min	1.0070	1.0117	1.0164	1.0263	9.8 min	1.0067	1.0118	1.0163	1.0267	9.5 min	1.0070	1.0115	1.0164	1.0261
Wavelet correlation																
	Level 1	14.2 s	0.4274	0.5544	0.6050	0.6628	13.1 s	0.4379	0.5628	0.6174	0.6704	12.5 s	0.4848	0.5762	0.6279	0.6697
30	Level 3	56.7 s	0.5732	0.7083	0.7868	0.8495	52.3 s	0.5738	0.7196	0.7939	0.8593	49.9 s	0.5908	0.7367	0.8096	0.8615
days	Level 5	3.8 min	0.7242	0.8312	0.8890	0.9355	3.5 min	0.7054	0.8454	0.9083	0.9443	3.3 min	0.7126	0.8643	0.9060	0.9457
before	Level 7	15.1 min	0.8089	0.8975	0.9443	0.9700	13.9 min	0.8018	0.9175	0.9530	0.9747	13.3 min	0.7944	0.9277	0.9555	0.9755
	Level 9	60.4 min	0.8264	0.9110	0.9536	0.9765	55.8 min	0.8257	0.9316	0.9623	0.9817	53.2 min	0.8285	0.9434	0.9674	0.9831
	Level 1	6.5 s	0.4718	0.5733	0.6190	0.6681	6.1 s	0.4892	0.5795	0.6236	0.6646	6.2 s	0.4774	0.5766	0.6251	0.6632
60	Level 3	26.1 s	0.5834	0.7288	0.7928	0.8686	24.2 s	0.5941	0.7370	0.8009	0.8555	24.7 s	0.5991	0.7397	0.8014	0.8653
days	Level 5	1.7 min	0.7124	0.8466	0.9045	0.9474	1.6 min	0.7452	0.8573	0.9029	0.9418	1.6 min	0.7207	0.8631	0.9039	0.9455
before	Level 7	7 min	0.8108	0.9207	0.9548	0.9769	6.5 min	0.8415	0.9294	0.9545	0.9737	6.6 min	0.8158	0.9287	0.9522	0.9760
	Level 9	27.8 min	0.8411	0.9416	0.9687	0.9845	25.9 min	0.8734	0.9489	0.9679	0.9824	26.3 min	0.8573	0.9510	0.9699	0.9844
	Level 1	4 s	0.4976	0.5849	0.6256	0.6668	3.7 s	0.5179	0.5880	0.6244	0.6639	3.7 s	0.5226	0.5959	0.6307	0.6662
3	Level 3	15.9 s	0.6325	0.7524	0.8065	0.8625	14.7 s	0.6332	0.7534	0.8047	0.8535	14.6 s	0.6542	0.7607	0.8123	0.8650
months	Level 5	1.1 min	0.7697	0.8663	0.9087	0.9450	58.7 s	0.7817	0.8693	0.9074	0.9405	58.5 s	0.7900	0.8744	0.9128	0.9453
before	Level 7	4.2 min	0.8621	0.9343	0.9573	0.9760	3.9 min	0.8758	0.9341	0.9564	0.9739	3.9 min	0.8798	0.9377	0.9584	0.9758
	Level 9	17 min	0.8930	0.9550	0.9724	0.9844	15.7 min	0.9079	0.9559	0.9712	0.9827	15.6 min	0.9148	0.9588	0.9736	0.9856
	Level 1	2.5 s	0.5677	0.6169	0.6431	0.6728	2.3 s	0.5691	0.6167	0.6425	0.6728	2.2 s	0.5727	0.6206	0.6448	0.6715
1	Level 3	10 s	0.7292	0.8001	0.8350	0.8725	9.2 s	0.7290	0.7988	0.8358	0.8733	8.9 s	0.7370	0.8019	0.8369	0.8740
Year	Level 5	39.8 s	0.8495	0.8985	0.9227	0.9483	36.9 s	0.8484	0.8998	0.9237	0.9496	35.4 s	0.8505	0.9004	0.9249	0.9491
before	Level 7	2.7 min	0.9211	0.9510	0.9642	0.9782	2.5 min	0.9198	0.9505	0.9643	0.9785	2.4 min	0.9231	0.9513	0.9648	0.9777
	Level 9	10.6 min	0.9454	0.9687	0.9776	0.9864	9.8 min	0.9435	0.9685	0.9777	0.9865	9.5 min	0.9476	0.9687	0.9780	0.9865

Notes: This table presents the average performance of EDM, EDU and EDZ median bid–ask covariance and median wavelet correlations within different timezones before maturity. Each group contains 10 Eurodollar contracts with the same delivery month in different years – such as EDM0, EDM1, . . . , EDM9, with average results for all four price levels.

I examine the short-term variance of high-frequency quotes. To do so, for each ED contract I rank trading days by the number of quotes and keep the top 10% most active days. Former analysis on the quoting variance revealed that the quoting variance suddenly drops approaching maturity (compared with Figures 2.5 and 2.6). Because the high trading volume period usually occurs within both an "equity-like" epoch and a "short-rate" epoch, therefore I only consider days up to two years prior to maturity and calculate the average variance ratios and the wavelet correlations for the high-speed quotes of each contract.

The results are exhibited in Table 2.9 which presents the top 10 ED contracts average high-speed quote variance and covariance based on the lowest timescales at level 1. The other 30 contract results are available in the online appendix. Consistent with my previous results, both the bid and ask variance ratios increase in moving to shorter timescales. The speed of quoting is definitely consistent with the existence of algorithmic traders; at time-scale level 1 and I observe quoting as fast as 280.5 ms and it is faster than 1 sec for all contracts.

Here again, a comparison with Hasbrouck [2014]'s results for the equity market is relevant to better understand the market of the Eurodollar Future. My results show that the high-speed quotes in the Eurodollar Future market have lower variance ratios and higher bid-ask correlations. The wavelet variance at timescale level 1 is one-half of the one reported by Hasbrouck for the US equity market. Hasbrouck reports that the variance ratio at the lowest timescale is about four times larger than that at the longest timescale in the equity markets, while my results suggest that the variance ratio at timescale level 1 is only about 1.5 times as large as the one observed at timescale level 9.

However, if I compare my high-frequency quoting results with Hasbrouck's highest quintile the differences are less pronounced. The results suggest that for all the

Table 2.9: The Average High-frequency Quoting Results from the Top 10 Contracts with the Most Active Days

	EDH0		EDH1		EDH9		EDM1		EDM9		EDU0		EDU9		EDZ0		EDZ8		EDZ9	
Top 10	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median
Ask variance																				
Level 1	341.7 ms	1.5236	354 ms	1.5187	322.1 ms	1.5273	363.3 ms	1.5251	288.7 ms	1.5257	391 ms	1.5210	280.5 ms	1.5321	357.9 ms	1.5198	387.5 ms	1.5249	296.6 ms	1.5273
Level 2	683.3 ms	1.2751	708.1 ms	1.2739	644.2 ms	1.2737	726.6 ms	1.2821	577.4 ms	1.2701	782 ms	1.2743	560.9 ms	1.2827	715.9 ms	1.2771	775 ms	1.2766	593.1 ms	1.2743
Level 3	1.4 s	1.1453	1.4 s	1.1574	1.3 s	1.1459	1.5 s	1.1626	1.2 s	1.1411	1.6 s	1.1496	1.1 s	1.1515	1.4 s	1.1588	1.6 s	1.1465	1.2 s	1.1492
Level 4	2.7 s	1.0815	2.8 s	1.0961	2.6 s	1.0814	2.9 s	1.1018	2.3 s	1.0776	3.1 s	1.0870	2.2 s	1.0843	2.9 s	1.1003	3.1 s	1.0817	2.4 s	1.0843
Level 5	5.5 s	1.0491	5.7 s	1.0643	5.2 s	1.0490	5.8 s	1.0672	4.6 s	1.0459	6.3 s	1.0544	4.5 s	1.0514	5.7 s	1.0659	6.2 s	1.0505	4.7 s	1.0519
Level 6	10.9 s	1.0324	11.3 s	1.0435	10.3 s	1.0327	11.6 s	1.0464	9.2 s	1.0301	12.5 s	1.0347	9 s	1.0341	11.5 s	1.0436	12.4 s	1.0340	9.5 s	1.0335
Level 7	21.9 s	1.0223	22.7 s	1.0290	20.6 s	1.0227	23.3 s	1.0303	18.5 s	1.0209	25 s	1.0233	18 s	1.0238	22.9 s	1.0288	24.8 s	1.0237	19 s	1.0233
Level 8	43.7 s	1.0161	45.3 s	1.0199	41.2 s	1.0171	46.5 s	1.0210	37 s	1.0157	50.1 s	1.0169	35.9 s	1.0174	45.8 s	1.0198	49.6 s	1.0172	38 s	1.0170
Level 9	1.5 min	1.0142	1.5 min	1.0166	1.4 min	1.0145	1.6 min	1.0169	1.2 min	1.0135	1.7 min	1.0148	1.2 min	1.0150	1.5 min	1.0165	1.7 min	1.0145	1.3 min	1.0147
Bid variance																				
Level 1	341.7 ms	1.5222	354 ms	1.5138	322.1 ms	1.5255	363.3 ms	1.5250	288.7 ms	1.5275	391 ms	1.5194	280.5 ms	1.5334	357.9 ms	1.5189	387.5 ms	1.5233	296.6 ms	1.5235
Level 2	683.3 ms	1.2748	708.1 ms	1.2725	644.2 ms	1.2744	726.6 ms	1.2843	577.4 ms	1.2699	782 ms	1.2713	560.9 ms	1.2827	715.9 ms	1.2758	775 ms	1.2757	593.1 ms	1.2745
Level 3	1.4 s	1.1455	1.4 s	1.1548	1.3 s	1.1451	1.5 s	1.1624	1.2 s	1.1414	1.6 s	1.1489	1.1 s	1.1505	1.4 s	1.1587	1.6 s	1.1469	1.2 s	1.1491
Level 4	2.7 s	1.0822	2.8 s	1.0959	2.6 s	1.0807	2.9 s	1.1008	2.3 s	1.0775	3.1 s	1.0873	2.2 s	1.0850	2.9 s	1.0979	3.1 s	1.0820	2.4 s	1.0848
Level 5	5.5 s	1.0494	5.7 s	1.0645	5.2 s	1.0492	5.8 s	1.0658	4.6 s	1.0460	6.3 s	1.0555	4.5 s	1.0518	5.7 s	1.0659	6.2 s	1.0505	4.7 s	1.0512
Level 6	10.9 s	1.0325	11.3 s	1.0438	10.3 s	1.0326	11.6 s	1.0446	9.2 s	1.0303	12.5 s	1.0366	9 s	1.0343	11.5 s	1.0440	12.4 s	1.0339	9.5 s	1.0338
Level 7	21.9 s	1.0229	22.7 s	1.0301	20.6 s	1.0226	23.3 s	1.0298	18.5 s	1.0209	25 s	1.0245	18 s	1.0240	22.9 s	1.0297	24.8 s	1.0237	19 s	1.0231
Level 8	43.7 s	1.0164	45.3 s	1.0205	41.2 s	1.0167	46.5 s	1.0212	37 s	1.0155	50.1 s	1.0179	35.9 s	1.0175	45.8 s	1.0205	49.6 s	1.0170	38 s	1.0170
Level 9	1.5 min	1.0143	1.5 min	1.0168	1.4 min	1.0144	1.6 min	1.0179	1.2 min	1.0135	1.7 min	1.0158	1.2 min	1.0151	1.5 min	1.0169	1.7 min	1.0144	1.3 min	1.0144
Bid–ask covariance																				
Level 1	341.7 ms	1.5205	354 ms	1.5133	322.1 ms	1.5263	363.3 ms	1.5213	288.7 ms	1.5246	391 ms	1.5182	280.5 ms	1.5319	357.9 ms	1.5175	387.5 ms	1.5220	296.6 ms	1.5227
Level 2	683.3 ms	1.2715	708.1 ms	1.2689	644.2 ms	1.2725	726.6 ms	1.2759	577.4 ms	1.2677	782 ms	1.2678	560.9 ms	1.2794	715.9 ms	1.2724	775 ms	1.2743	593.1 ms	1.2700
Level 3	1.4 s	1.1390	1.4 s	1.1468	1.3 s	1.1425	1.5 s	1.1523	1.2 s	1.1370	1.6 s	1.1411	1.1 s	1.1458	1.4 s	1.1528	1.6 s	1.1429	1.2 s	1.1424
Level 4	2.7 s	1.0739	2.8 s	1.0874	2.6 s	1.0753	2.9 s	1.0908	2.3 s	1.0733	3.1 s	1.0765	2.2 s	1.0784	2.9 s	1.0919	3.1 s	1.0771	2.4 s	1.0772
Level 5	5.5 s	1.0400	5.7 s	1.0543	5.2 s	1.0425	5.8 s	1.0560	4.6 s	1.0396	6.3 s	1.0437	4.5 s	1.0439	5.7 s	1.0577	6.2 s	1.0445	4.7 s	1.0419
Level 6	10.9 s	1.0226	11.3 s	1.0343	10.3 s	1.0259	11.6 s	1.0333	9.2 s	1.0231	12.5 s	1.0254	9 s	1.0252	11.5 s	1.0367	12.4 s	1.0279	9.5 s	1.0242
Level 7	21.9 s	1.0128	22.7 s	1.0207	20.6 s	1.0160	23.3 s	1.0203	18.5 s	1.0143	25 s	1.0147	18 s	1.0153	22.9 s	1.0229	24.8 s	1.0175	19 s	1.0143
Level 8	43.7 s	1.0075	45.3 s	1.0127	41.2 s	1.0102	46.5 s	1.0122	37 s	1.0090	50.1 s	1.0090	35.9 s	1.0094	45.8 s	1.0139	49.6 s	1.0111	38 s	1.0087
Level 9	1.5 min	1.0051	1.5 min	1.0086	1.4 min	1.0072	1.6 min	1.0080	1.2 min	1.0064	1.7 min	1.0062	1.2 min	1.0067	1.5 min	1.0095	1.7 min	1.0078	1.3 min	1.0061
Wavelet correlation																				
Level 1	341.7 ms	0.6569	354 ms	0.6606	322.1 ms	0.6555	363.3 ms	0.6557	288.7 ms	0.6547	391 ms	0.6582	280.5 ms	0.6522	357.9 ms	0.6584	387.5 ms	0.6565	296.6 ms	0.6564
Level 2	683.3 ms	0.7844	708.1 ms	0.7859	644.2 ms	0.7847	726.6 ms	0.7786	577.4 ms	0.7874	782 ms	0.7866	560.9 ms	0.7796	715.9 ms	0.7839	775 ms	0.7839	593.1 ms	0.7846
Level 3	1.4 s	0.8730	1.4 s	0.8659	1.3 s	0.8733	1.5 s	0.8603	1.2 s	0.8761	1.6 s	0.8704	1.1 s	0.8692	1.4 s	0.8630	1.6 s	0.8719	1.2 s	0.8703
Level 4	2.7 s	0.9240	2.8 s	0.9125	2.6 s	0.9253	2.9 s	0.9084	2.3 s	0.9281	3.1 s	0.9198	2.2 s	0.9216	2.9 s	0.9108	3.1 s	0.9242	2.4 s	0.9219
Level 5	5.5 s	0.9529	5.7 s	0.9394	5.2 s	0.9531	5.8 s	0.9383	4.6 s	0.9560	6.3 s	0.9474	4.5 s	0.9507	5.7 s	0.9382	6.2 s	0.9520	4.7 s	0.9513
Level 6	10.9 s	0.9685	11.3 s	0.9580	10.3 s	0.9684	11.6 s	0.9573	9.2 s	0.9706	12.5 s	0.9647	9 s	0.9668	11.5 s	0.9578	12.4 s	0.9672	9.5 s	0.9673
Level 7	21.9 s	0.9776	22.7 s	0.9708	20.6 s	0.9779	23.3 s	0.9711	18.5 s	0.9795	25 s	0.9760	18 s	0.9766	22.9 s	0.9711	24.8 s	0.9769	19 s	0.9774
Level 8	43.7 s	0.9839	45.3 s	0.9799	41.2 s	0.9835	46.5 s	0.9792	37 s	0.9848	50.1 s	0.9824	35.9 s	0.9828	45.8 s	0.9799	49.6 s	0.9833	38 s	0.9833
Level 9	1.5 min	0.9859	1.5 min	0.9835	1.4 min	0.9858	1.6 min	0.9824	1.2 min	0.9866	1.7 min	0.9844	1.2 min	0.9851	1.5 min	0.9833	1.7 min	0.9858	1.3 min	0.9858

Notes: I extract those days within two years prior to the maturity of each ED contract. This table only presents high-frequency quoting results of 10 ED contracts with the lowest level 1 quote timescale. The other 30 contract results are available in the online appendix.

Table 2.10: High-frequency Quoting Results of the Most Active Day for Each Contract

	EDH0		EDH1		EDH9		EDM0		EDM1		EDM9		EDU9		EDZ0		EDZ8		EDZ9	
Most active day	**June 5, 2009**		**May 7, 2010**		**September 17, 2008**		**May 7, 2010**		**January 22, 2008**		**September 16, 2008**		**March 14, 2008**		**September 16, 2008**		**May 7, 2010**		**June 5, 2009**	
	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median	Time scale	Median
Ask variance																				
Level 1	235.3 ms	1.5968	255.8 ms	1.5614	217.5 ms	1.5009	266.5 ms	1.5888	252.4 ms	1.5652	223.3 ms	1.4959	194.4 ms	1.6167	262.2 ms	1.5374	251.5 ms	1.5277	109.8 ms	1.5951
Level 2	470.5 ms	1.3193	511.6 ms	1.3584	435 ms	1.2233	533 ms	1.3139	504.8 ms	1.3555	446.5 ms	1.2419	388.9 ms	1.3026	524.5 ms	1.3332	502.9 ms	1.2357	219.7 ms	1.3997
Level 3	941.1 ms	1.1619	1 s	1.2242	870 ms	1.1062	1.1 s	1.1612	1 s	1.2139	893.1 ms	1.1146	777.7 ms	1.1499	1 s	1.2056	1 s	1.1115	439.3 ms	1.2944
Level 4	1.9 s	1.0867	2 s	1.1234	1.7 s	1.0534	2.1 s	1.0856	2 s	1.1173	1.8 s	1.0535	1.6 s	1.0803	2.1 s	1.1104	2 s	1.0561	878.7 ms	1.2214
Level 5	3.8 s	1.0516	4.1 s	1.0689	3.5 s	1.0290	4.3 s	1.0488	4 s	1.0665	3.6 s	1.0279	3.1 s	1.0445	4.2 s	1.0601	4 s	1.0309	1.8 s	1.1634
Level 6	7.5 s	1.0344	8.2 s	1.0416	7 s	1.0170	8.5 s	1.0325	8.1 s	1.0389	7.1 s	1.0170	6.2 s	1.0263	8.4 s	1.0353	8 s	1.0176	3.5 s	1.1162
Level 7	15.1 s	1.0256	16.4 s	1.0254	13.9 s	1.0117	17.1 s	1.0236	16.2 s	1.0230	14.3 s	1.0121	12.4 s	1.0179	16.8 s	1.0212	16.1 s	1.0118	7 s	1.0840
Level 8	30.1 s	1.0199	32.7 s	1.0166	27.8 s	1.0093	34.1 s	1.0181	32.3 s	1.0144	28.6 s	1.0093	24.9 s	1.0136	33.6 s	1.0139	32.2 s	1.0092	14.1 s	1.0625
Level 9	1 min	1.0173	1.1 min	1.0136	55.7 s	1.0084	1.1 min	1.0158	1.1 min	1.0114	57.2 s	1.0081	49.8 s	1.0122	1.1 min	1.0112	1.1 min	1.0084	28.1 s	1.0548
Bid variance																				
Level 1	235.3 ms	1.5984	255.8 ms	1.5603	217.5 ms	1.4994	266.5 ms	1.5911	252.4 ms	1.5702	223.3 ms	1.4957	194.4 ms	1.6189	262.2 ms	1.5389	251.5 ms	1.5282	109.8 ms	1.6005
Level 2	470.5 ms	1.3199	511.6 ms	1.3627	435 ms	1.2213	533 ms	1.3145	504.8 ms	1.3597	446.5 ms	1.2415	388.9 ms	1.3036	524.5 ms	1.3354	502.9 ms	1.2375	219.7 ms	1.4003
Level 3	941.1 ms	1.1635	1 s	1.2287	870 ms	1.1054	1.1 s	1.1591	1 s	1.2213	893.1 ms	1.1145	777.7 ms	1.1511	1 s	1.2108	1 s	1.1120	439.3 ms	1.2964
Level 4	1.9 s	1.0895	2 s	1.1286	1.7 s	1.0527	2.1 s	1.0849	2 s	1.1245	1.8 s	1.0532	1.6 s	1.0812	2.1 s	1.1151	2 s	1.0561	878.7 ms	1.2384
Level 5	3.8 s	1.0528	4.1 s	1.0722	3.5 s	1.0284	4.3 s	1.0499	4 s	1.0695	3.6 s	1.0276	3.1 s	1.0441	4.2 s	1.0628	4 s	1.0307	1.8 s	1.1788
Level 6	7.5 s	1.0350	8.2 s	1.0436	7 s	1.0165	8.5 s	1.0330	8.1 s	1.0405	7.1 s	1.0168	6.2 s	1.0260	8.4 s	1.0364	8 s	1.0175	3.5 s	1.1227
Level 7	15.1 s	1.0254	16.4 s	1.0262	13.9 s	1.0112	17.1 s	1.0232	16.2 s	1.0237	14.3 s	1.0117	12.4 s	1.0178	16.8 s	1.0214	16.1 s	1.0116	7 s	1.0829
Level 8	30.1 s	1.0191	32.7 s	1.0168	27.8 s	1.0090	34.1 s	1.0178	32.3 s	1.0149	28.6 s	1.0091	24.9 s	1.0137	33.6 s	1.0137	32.2 s	1.0091	14.1 s	1.0586
Level 9	1 min	1.0169	1.1 min	1.0135	55.7 s	1.0084	1.1 min	1.0157	1.1 min	1.0116	57.2 s	1.0080	49.8 s	1.0122	1.1 min	1.0111	1.1 min	1.0085	28.1 s	1.0478
Bid–ask covariance																				
Level 1	235.3 ms	1.5959	255.8 ms	1.5593	217.5 ms	1.4999	266.5 ms	1.5887	252.4 ms	1.5658	223.3 ms	1.4956	194.4 ms	1.6173	262.2 ms	1.5374	251.5 ms	1.5278	109.8 ms	1.5861
Level 2	470.5 ms	1.3159	511.6 ms	1.3572	435 ms	1.2217	533 ms	1.3111	504.8 ms	1.3542	446.5 ms	1.2411	388.9 ms	1.3020	524.5 ms	1.3329	502.9 ms	1.2360	219.7 ms	1.3784
Level 3	941.1 ms	1.1568	1 s	1.2218	870 ms	1.1049	1.1 s	1.1549	1 s	1.2124	893.1 ms	1.1135	777.7 ms	1.1487	1 s	1.2058	1 s	1.1108	439.3 ms	1.2708
Level 4	1.9 s	1.0797	2 s	1.1196	1.7 s	1.0517	2.1 s	1.0778	2 s	1.1145	1.8 s	1.0517	1.6 s	1.0780	2.1 s	1.1096	2 s	1.0548	878.7 ms	1.2046
Level 5	3.8 s	1.0418	4.1 s	1.0633	3.5 s	1.0268	4.3 s	1.0400	4 s	1.0613	3.6 s	1.0256	3.1 s	1.0410	4.2 s	1.0576	4 s	1.0291	1.8 s	1.1451
Level 6	7.5 s	1.0235	8.2 s	1.0355	7 s	1.0144	8.5 s	1.0221	8.1 s	1.0331	7.1 s	1.0142	6.2 s	1.0223	8.4 s	1.0318	8 s	1.0154	3.5 s	1.0944
Level 7	15.1 s	1.0145	16.4 s	1.0194	13.9 s	1.0087	17.1 s	1.0129	16.2 s	1.0174	14.3 s	1.0091	12.4 s	1.0138	16.8 s	1.0175	16.1 s	1.0092	7 s	1.0606
Level 8	30.1 s	1.0091	32.7 s	1.0110	27.8 s	1.0061	34.1 s	1.0079	32.3 s	1.0094	28.6 s	1.0062	24.9 s	1.0092	33.6 s	1.0101	32.2 s	1.0063	14.1 s	1.0383
Level 9	1 min	1.0063	1.1 min	1.0074	55.7 s	1.0047	1.1 min	1.0053	1.1 min	1.0061	57.2 s	1.0045	49.8 s	1.0070	1.1 min	1.0071	1.1 min	1.0047	28.1 s	1.0265
Wavelet correlation																				
Level 1	235.3 ms	0.6256	255.8 ms	0.6409	217.5 ms	0.6669	266.5 ms	0.6285	252.4 ms	0.6369	223.3 ms	0.6686	194.4 ms	0.6177	262.2 ms	0.6498	251.5 ms	0.6544	109.8 ms	0.6248
Level 2	470.5 ms	0.7577	511.6 ms	0.7339	435 ms	0.8188	533 ms	0.7607	504.8 ms	0.7355	446.5 ms	0.8054	388.9 ms	0.7671	524.5 ms	0.7488	502.9 ms	0.8081	219.7 ms	0.7142
Level 3	941.1 ms	0.8595	1 s	0.8139	870 ms	0.9046	1.1 s	0.8627	1 s	0.8188	893.1 ms	0.8972	777.7 ms	0.8687	1 s	0.8259	1 s	0.8993	439.3 ms	0.7714
Level 4	1.9 s	0.9179	2 s	0.8860	1.7 s	0.9499	2.1 s	0.9217	2 s	0.8893	1.8 s	0.9495	1.6 s	0.9249	2.1 s	0.8968	2 s	0.9469	878.7 ms	0.8075
Level 5	3.8 s	0.9498	4.1 s	0.9327	3.5 s	0.9724	4.3 s	0.9524	4 s	0.9350	3.6 s	0.9731	3.1 s	0.9578	4.2 s	0.9409	4 s	0.9702	1.8 s	0.8483
Level 6	7.5 s	0.9662	8.2 s	0.9582	7 s	0.9837	8.5 s	0.9681	8.1 s	0.9611	7.1 s	0.9835	6.2 s	0.9747	8.4 s	0.9649	8 s	0.9828	3.5 s	0.8907
Level 7	15.1 s	0.9753	16.4 s	0.9745	13.9 s	0.9890	17.1 s	0.9774	16.2 s	0.9769	14.3 s	0.9884	12.4 s	0.9825	16.8 s	0.9790	16.1 s	0.9885	7 s	0.9234
Level 8	30.1 s	0.9812	32.7 s	0.9835	27.8 s	0.9911	34.1 s	0.9825	32.3 s	0.9854	28.6 s	0.9910	24.9 s	0.9865	33.6 s	0.9865	32.2 s	0.9910	14.1 s	0.9447
Level 9	1 min	0.9834	1.1 min	0.9867	55.7 s	0.9917	1.1 min	0.9845	1.1 min	0.9886	57.2 s	0.9921	49.8 s	0.9880	1.1 min	0.9891	1.1 min	0.9916	28.1 s	0.9543

Notes: I extract each highest speed quoting day within two years prior to the maturity for each ED contract. This table only presents high-frequency quoting results of 10 ED contracts with the lowest level 1 quote timescale. The other 30 contract results are available in the online appendix.

contracts the bid and ask variances are around 1.5 at timescale level one, and reduce to near 1 at the longest time-scale. Hasbrouck's equity results range from 1.37 in the shortest timescales to 1 when moving to the longest timescales. The bid-ask co-movements can be assessed by wavelet bid–ask covariances and correlations at different timescales. My wavelet bid–ask correlations of high-speed quotes are also similar to the correlations in the US equity market reported by Hasbrouck. With these results in hand, I consider the period between two years before maturity and one week to maturity to be an "equity-like" epoch in that the Eurodollar Future market exhibits similar characteristics than the ones observed in the American equity market.

For robustness, this chapter also considers only the most active days for each contract, instead of the top 10% active days. Table 2.10 presents the variance ratio of the highest daily number of bids for 10 ED contracts with the lowest level 1 quote timescale. The other 30 ED contract-code results are available in the online appendix. Timescales of the most active days are much smaller compared with Table 2.9. The shortest timescale is 109.8 milliseconds at level one of EDZ9. Wavelet variance ratios and correlations are unity with the top 10% days' results. Both bid variance ratio and asks variance ratio decline from around 1.5 at level one to about 1 at the longest timescale. Wavelet bid-ask correlations of the most active days are higher than 0.6 at the shortest timescale and close to 1 at the longest timescale. Overall, the results of the most active days in Table 2.10 are materially similar although in general, the timescales are much smaller than the timescales of the top 10% active days.

The results I have discussed so far allow us to characterize the trading life-cycle of the Eurodollar Future that I present in Figure 2.1. My initial descriptive statistics suggest that there is not much action in terms of quotes neither in terms of trades

during the first two years of life of the Eurodollar Future. Indeed, during this initial period both transactions and quoting are infrequent. I label this period as a "thin trading" epoch. This period is then followed by an "equity-like"epoch when high frequency quoting is very intense. This period is characterized by low variance ratios in both sides of the book and low correlations between asks and bids. Finally, as the contract get closer to maturity this frenetic activity declines. I label this period "short rate" epoch, characterized by decreasing quoting frequency, high volumes and high variance ratios.

2.5 Chapter Summary

I have provided an extensive empirical analysis of the microstructure of bid and ask volatility on the Eurodollar Futures market during the period 1996–2014 utilizing a Haar wavelet multi-resolution analysis on every inside-quote update across the whole market. Whilst I find equity-like variance ratios at certain points in the maturity life-cycle of the money market futures, this effect varies considerably with maturity. I have identified several specific epochs, the initial seven years or so of the contract (the furthest tenors in the term structure) is characterized by little no activity; however, between two and one years from maturity, for which the trading activity rises substantially.

However, even when the speed of update approaches 100 milliseconds, my variance ratios although greater than one are not at levels found in the equity literature. In contrast, though wavelet correlations between the bid and the ask prices do drop considerably and this is in keeping with the equity literature, indicating that the order books in these markets have a great deal of noise, very likely from active HFT automated trading. Unlike in an equity market where multiple overlapping

exchanges create bottlenecks, the CME is centrally cleared with microsecond (although in practice 0.1 millisecond) time stamping and latency. So the playing field is relatively even, from the market design point of view, the fact that I do not see exploding variance ratios during the high trading epoch could be useful in any potential redesign of the national market system. It is worth remembering that at any given time, eight of these contracts will be within the equity like-epoch and one quarterly and two monthly contracts will be within the short rate epoch.

So it appears that HFT shifts trading once the contracts pass a critical time-to-maturity. Identification of this critical time, maybe of help to those seeking to avoid direct trading against HFTs, however as I discussed previously, those hedging swaps with IMM dates are substantively time constrained as the purpose of the transaction is hedging specifically date floating rate risk. Whilst my data contains pit-trades, which are by construction, not HFT, the low number of trades renders analysis by wavelet-based variance ratio analysis impractical.

I do, however, find that very near to maturity, as the future mimics the short rate, the variance ratios drop dramatically. This epoch is usually as the contract approaches one or two weeks from maturity. However, the speed of updating in this period is not high and at high frequencies whilst the wavelet correlations decrease it is not to the extent found in the preceding equity-like period. Taken together this results paint an interesting picture that should provide a helpful guide to practitioners and academics seeking to identify the time scale at which microstructure effects begin to inundate themselves on the price process, for a limit order book in a centrally cleared market. Realized volatility analysis at the sub-second level is increasingly common for estimating the ex-post quadratic variation in prices and the bid-ask spread plays a central role in providing a proxy for market illiquidity. The results indicate that realized volatility estimation for

the long maturity futures will simply not be possible as trading activity is so low. In contrast, for one to two years from maturity sub-second measurements are possible and, from my results, less noisy that those found in the equity markets. Given the scope of this study, a large variety of options exists for summarizing the variance and covariance ratios. An online appendix provides a number of different slices across the data, by contract, maturity trading frequency, calendar time and conducts a comprehensive series of robustness checks for the interested reader.

Chapter 3

High Speed Trading Behavior and Market Liquidity: Evidence From Eurodollar Limited Order Book Data

Using the Eurodollar futures market as a case study, I have constructed a unique dataset consisting of both the inside quotes (best-bid, best-ask) – every update from 1996 to 2014 – and the entire activity within the limit order book – from 2008 to 2014. This allows me to classify, ex-post, the realized fraction of high frequency trading activity within the market. I find that the most active level of the order book for the overall volume of quotes and in particular high speed quoting is level two. Indicating that studies on this type of contracts that ignore quoting outside the best-bid best-ask are possibly missing some interesting features of the market. I demonstrate theoretically that a non-linear relationship between the proportion of informed trading and the market quality and critical HFT saturation levels would be anticipated. My results paint a fascinating picture of the changing behavior of traders in these very actively traded contracts and the market liquidity over the texrm structure of the Eurodollar contract.

3.1 Introduction

Empirical research on the microstructure of interest rate derivatives is a relatively overlooked area of financial economics. The futures contract on the 3-month Eurodollar deposit is one of the most actively traded in global financial markets. This appears to be an anomaly as reliable estimates indicate that the notional dollar amount outstanding on these contracts potentially makes this one of the world's largest financial markets.

In this chapter, I conduct an extensive analysis of the three-month Eurodollar futures market and in particular the short-term, intraday, volatility and liquidity of over the tenor of the futures contract. I will demonstrate that high frequency trading activity has a relatively stable term structure over my sample period. Specifically, I will make use of both the inside quotes and all of the other quotes from the limit order book, and classify trading activity by speed of update of quotes within the market. My dataset includes the number of active accounts connected to the market and this allows me to instrument for the likelihood of simultaneous execution resulting in the accidental false-positive identification of high frequency quoting in a busy market. Indeed, the fractions extracted from the order-book are qualitative close to those computed by the Chicago Mercantile Exchange in a similar study, where participants were required to declare the connection type to the exchange, information not made publicly available to researchers.

My dataset is available to any researchers to buy, however, to structure it in a way that is useful for econometric analysis requires substantial pre-processing. Each microsecond time stamped update to the order book contains a snapshot of the current market across all the actively traded levels sorted by price. Typically

a single contract during its actively traded epoch (see Figure 2.1 for a graphical summary) will have several million rows of data each day by five to ten quoting levels (moving out from the inside quotes at level one). For each level at each time stamp, there are three pieces of information, the price of the quote, the volume of contracts bid or asked and the number of active accounts with open quotes. CME group has access to the list of individual accounts and this is not made available publicly to researchers. This means I cannot place a specific quantity on the number of actual trading accounts in the market, only provide a notional maximum assuming that at any give time stamp that the same accounts are not quoting across a variety of levels on both sides of the order book. However, even on a very busy day the maximum number of active accounts does not appear to exceed 100, and it seems more likely to be less than 30. Given that the median update time of the order book is between 3 and 4 milliseconds, it appears unlikely that the realized daily fraction of high frequency updates is inflated by simultaneous execution being miss-classified as high frequency trading. Furthermore, the CME's own analysis, see CME Group [2010], of the fraction of algorithmic trading accounts active during a day is very much in keeping with unconditional averages I compute herein, but the classification of high frequency updates as those occurring at a speed substantially beyond normal human reaction times (200 ms). I further conduct a sensitivity analysis to demonstrate that the realized fraction of detected high frequency traders within the market does not change substantially as I move from a threshold of 200 ms to 25 ms.

Indeed, Keynes predicted that in futures markets with hedgers and speculators with different market participation requirements indicates that a premium should be paid by seekers of protection from future variation in the spot price to those providing that protection. However, the equilibria in such markets will of course depend on the proportion of hedgers and speculators and the degree of technolog-

ical differentiation between them. After extraction of the realized high frequency trading activity proxy, I therefore eschew linear regression for a semi-parametric estimator when attempting to determine the relative execution risk as a function of potential algorithmic trading in the next chapters, which I presume to be a technology of the speculator.

Imposing the restriction that algorithm based high frequency trading is a speculator technology permits a simpler interpretation of my results; however, relaxing this assumption does not diminish the ability of the AT regressions to identify the mechanisms that determine the impact of algorithmic high speed trading on execution risk and the quality of the market. However, if hedgers are also engaged in AT trading, the interpretation of non-linear interactions between the fraction of algorithmic-based traders in the market is inherently more complex. For instance, I postulate that there will be a term structure to AT trading based around the fact that IMM dates play an important role in the institutional provision of interest rate swaps (IRSs). From the perspective of AT as a pure speculation instrument, I would anticipate that the algorithms would deteriorate the quality of the market up to a point then effectively cancel themselves out as the speculators formed the majority of market participants. However, if both hedgers and speculators used use AT, then the interaction would be less clear as the key driver would be the technology gradient between the market participants and this would not provide for specific predictions.

This chapter introduces an asymmetric information based theoretical model for futures trading to predict the effects of informed traders on the term structure of Eurodollar future contracts. Following the approach of Hendershott et al. [2011] utilized in the equity market, I measure a battery of market liquidity spreads on every level of the Eurodollar future limit order book data from 2008

to 2014, including bid-ask spreads, quoted half spreads, quoted depth, effective spreads, realized spreads and adverse selection indicators. I also measure the proportion of high frequency algorithmic trading as the AT proxy, based on the quotes messaging updates under a certain time thresholds (from 200 milliseconds to 25 milliseconds). To my knowledge, this is the only study of its type to effectively cover the entire population of quote updates and the only one to cover the microstructure of this market. I will demonstrate that: a) theoretically, the fraction of informed trading in the market has a non-linear marginal effect on the market quality. b) not including quoting activity over and above the inside quotes loses a great deal of information. Indeed, I show that for the majority of the term structure of the contracts the level two quotes are by far the most active.

The remainder of this chapter is organized as follows: Section 3.2 outlines the theoretical development with an asymmetric information based model of futures trading, and Section 3.3 summarizes related literature. Section 3.4 presents the empirical methodologies to calculate a set of market liquidity indicators and high frequency algorithmic trading proxies. Section 3.5 and Section 3.6 describe my Eurodollar future dataset and provide some analysis from the market liquidity and the high-speed algorithmic trading proxies from the Eurodollar limited order book. Section 3.7 provides some summary comments.

3.2 Theoretical Model for Futures Trading

3.2.1 A Theoretical Model of Advantaged Trader Saturation

My major empirical observation is that the fraction of informed trading has a non-linear relationship with the execution risk and other market quality indicators, whereby the impact of low numbers of traders with significant trading advantages can have a deleterious effect on market quality, but that this patter reverses as more traders enter the market. We can see that this type of effect is predicted in a simple linear noisy rational expectations model in the spirit of Admati and Pfleiderer [1988] and Admati and Pfleiderer [1989].

Let $\tilde{f}(t) = [\tilde{f}_S(t), \tilde{f}_L(t)]'$ for $f(t) \in \mathbb{R}^2$ be the log prices of pair of futures contracts. I presume that neither contract is maturing that day, but that the reference rate is revealed such that the final prices for the day at mark to market $\tilde{\delta}(t) - f(T)$ is the variance between the fair valuation and the marked to market futures prices. Indeed, in the Eurodollar futures market the reference yield curve from the dollar LIBOR rates is revealed several hours (recall this is a 24hr market) prior to the daily settlements (11 am CET, with settlement at 20:00 CET).

I will restrict my theoretical development to two contracts simultaneously traded in the market although my results are generalizable and analytically tractable for n_f futures. I will presume that the maturity date of $\tilde{f}_S(t)$ is $\bar{T}_S$ and $\tilde{f}_L(t)$ is $\bar{T}_L$ where $\bar{T}_L > \bar{T}_S$ and $T_{i\in\{L,S\}} > T$, therefore neither future is maturing within the days trading and hence there is no final settlement. I can think of $\tilde{f}_S(t)$ being the short maturity future and $\tilde{f}_L(t)$ being the long maturity future. I presume that there are $\tilde{N}$ traders in the market and this is split between $N_A$and N_B, such that

$\tilde{N} = N_A + N_B$. I model a days trading in a future such that the mark to market valuation is denoted $\tilde{\delta}(T)$. I will presume that trading takes place continuously over a days trading, such that:

$$\mathrm{d}\tilde{\delta}(t) = \tilde{\Sigma}^{1/2} d\tilde{W}(t) \tag{3.1}$$

where $\tilde{W}(t)$ is a two-dimensional brownian motion and subsequently $\delta(t + \Delta t) - \delta(t) \sim \mathcal{N}(0, \Delta t \tilde{\Sigma})$ over the interval $0 \leq t \leq T$. I presume that a maturity effect exists therefore setting $\Sigma = [\sigma_{ij}]_{ij \in \{S,L\}}$, I impose that $\sigma_{SS} > \sigma_{LL}$. Under the fair valuation assumption, the final true value of the futures $V(T)$ is known and agreed on with certainty by all participants.

I denote with the subscript B uninformed traders, who submit a random aggregate order flow of vector $\tilde{d}$ for each of the futures over the interval $0, T$, this order flow is presumed to be zero centered $\mathbb{E}_0[\tilde{d}] = 0$ covariance matrix $\mathbb{E}_0[\tilde{d}\tilde{d}'] = N_B \tilde{\Psi}$. In the Kyle [1985], Glosten and Milgrom [1985] and Admati [1985] approach this is the noise or liquidity traders providing aggregate liquidity to the market. I can think of this as institutional traders are forced to create specific positions in the market. This order flow is presumed to be evenly distributed over the interval $0, T$.

The group of traders denoted by the subscript A is algorithmic high speed traders who trade off the order-flow information. It is at this point I depart from the standard Kyle [1985], Glosten and Milgrom [1985] and Admati [1985] approach. In a futures market, the mark to market prices are assumed to be set at the end of a days trading denoted T. Traders denoted A are presumed to have a technical advantage over the institutional traders as they can track the direction of the order flow and as such the have an unbiased but noisy valuation of the terminal

valuation $\delta(T)$, whereby the $n \in N_A$ receives a signal $\delta_A(T)$ such that

$$\tilde{\delta}_A(T) = \tilde{\delta}(T) + \tilde{\zeta} + \tilde{\xi} \tag{3.2}$$

where $\tilde{\zeta} \in \mathcal{N}(0, \tilde{\Gamma})$ and $\tilde{\xi} \in \mathcal{N}(0, \tilde{\Phi})$ are the global and trader specific noise vectors respectively. For tractability, I do not model their instantaneous profit function, but I can write down their objective in the form of their integrated profit function, let $\tilde{a}_n$ be the aggregate order flow of the $n \in N_A$ trader

$$\tilde{a}_n^* = \arg\max_{\tilde{a}_n} \mathbb{E}[(\tilde{\delta}(T) - f(T))'\tilde{a}_n] \tag{3.3}$$

I presume that there is a market clearing mechanism within the futures market that matches order flow. Let $\tilde{e} = \sum_{n=1}^{N_A} \tilde{a}_n^* + \tilde{d}$ be the aggregate order flow across both futures contracts. In a centrally cleared trading platform such as CME, the clearing house matches and clears all available trades and then places the remainder in the limit order book. The mechanism can be thought of as a price matching model whereby the objective is to promote the efficient clearing of information such that $\tilde{\delta}(t) - f(T) \to 0$. Therefore, the market clearing pricing rule is founded on the principal that $\mathbb{E}_0[f(T)] = \tilde{\delta}(t)$.

It is useful at this juncture to specify a quick notational convention I specify a positive semi-definite matrix $\tilde{\Theta} \in \mathbb{C}^{2\times 2}$ and subsequently let $\tilde{\Theta} \in \{\tilde{\Sigma}, \tilde{\Gamma}, \tilde{\Phi}, N_B\tilde{\Psi}\}$ represent the structural parameter matrices. I denote the individual elements as $\tilde{\Theta} = [\tilde{\theta}_{ij}]_{ij\in\{S,L\}}$. I decompose the positive definite matrix $\tilde{\Theta} = \Theta\Theta'$, where Θ is lower triangular with elements $\Theta = [\theta_{ij}]$ and $\theta_{i>j} = 0$, the elements of this matrix are labelled vech$\Theta = [\theta_k]'_{k\in\{1,\dots,3\}}$. Recalling that vech$\tilde{\Theta} = [\theta_{SS}, \theta_{SL}, \theta_{LL}]'$. Let $\tilde{\theta}^{1/2}$ be the Cholesky factorization operator and eig be the eigenvalue operator such that eig : $\mathbf{A} \to \mathbf{a}$ reports the vector of eigenvalues for a square matrix $\mathbf{A}$.

The model can be expressed as a simple one period Kalman filter assuming that the matrices $\tilde{\Theta} \in \{\tilde{\Sigma}, \tilde{\Gamma}, \tilde{\Phi}, N_B\tilde{\Psi}\}$ and the scalar N_A are known a-priori by all participants.[1] This leads me to my first result, the market equilibrium under rational expectations.

Theorem 1. *A Noisy Linear Rational Expectation Market Equilibrium. Let the observed price of the futures contracts be $f(t) = \tilde{\delta}(t) + \tilde{\Lambda}\tilde{e}$, where $\tilde{\Lambda}$ is the multivariate equivalent of "Kyle's lambda" and $\tilde{e}$ is net order flow, with the algorithmic traders linear noisy rational expectations equilibrium is given by $\tilde{a}_n^* = \tilde{B}\tilde{\delta}(t)$. The market clearing equilibrium under Noisy Linear Rational Expectations (NLRE) is defined by the following equations of state:*

$$\tilde{\Lambda} = \sqrt{N_A}(N_B\tilde{\Psi})^{-1/2}\tilde{M}^{-1/2}(N_B\tilde{\Psi})^{-1/2}, \quad \tilde{B} = \tilde{\Lambda}^{-1}\tilde{J}^{-1}\tilde{\Sigma}^{-1} \tag{3.4}$$

$$\tilde{M} = \sqrt{N_B}\tilde{\Psi}^{1/2}\tilde{J}^{-1}\tilde{\Sigma}_\xi\tilde{J}'^{-1}\sqrt{N_B}\tilde{\Psi}^{1/2}, \quad \tilde{\Sigma}_\xi = \tilde{\Sigma}(\tilde{\Sigma}+\tilde{\Gamma}+\tilde{\Phi})^{-1}\tilde{\Sigma}' \tag{3.5}$$

$$\tilde{J} = 2I + (N_A - 1)\tilde{\Sigma}(\tilde{\Sigma}+\tilde{\Gamma}+\tilde{\Phi})^{-1}(\tilde{\Sigma}+\tilde{\Gamma})\tilde{\Sigma}^{-1} \tag{3.6}$$

The proof proceeds in effectively the same manner as the proof in Watanabe [2008], pages 33-39, however I generalize it to the case where the $\tilde{N} = [N_i]$ is a diagonal matrix of informed traders for each asset. I augment the details of the

[1] A useful result to note is that if $\text{vech}\Theta \in \mathbb{R}^3_+$ and $\theta_2 < \theta_1\theta_3$ then the Cholesky factors recover the elements of Θ and the eigenvalues of the matrix $\tilde{\Theta}$ are analytic and greater than zero.

$$\tilde{\Theta}^{1/2} = \begin{pmatrix} \sqrt{\theta_1^2} & \frac{\theta_1\theta_2|\theta_1|^{-1}}{\sqrt{-\theta_2^2+\theta_2^2+\theta_3^2}} \\ 0 & \end{pmatrix} \equiv \Theta \quad \text{and} \quad \tilde{\Theta}^{-1/2} \equiv (\tilde{\Theta}^{-1})^{1/2} = \begin{pmatrix} \frac{\sqrt{\theta_2^2+\theta_3^2}}{\theta_1\theta_3} & -\frac{\theta_2}{\theta_3\sqrt{\theta_2^2+\theta_3^2}} \\ 0 & \frac{1}{\sqrt{\theta_2^2+\theta_3^2}} \end{pmatrix}$$

and

$$\text{eig}\,\tilde{\Theta} = \frac{1}{2}\begin{pmatrix} \theta_1^2+\theta_2^2+\theta_3^2 - \sqrt{(\theta_1^2+\theta_2^2+\theta_3^2)^2 - 4\theta_1^2\theta_3^2} \\ \theta_1^2+\theta_2^2+\theta_3^2 + \sqrt{(\theta_1^2+\theta_2^2+\theta_3^2)^2 - 4\theta_1^2\theta_3^2} \end{pmatrix}$$

it is useful to note that the presumption that $\theta_2 > 0$ is not needed to recover an analytically tractable result; however, I commonly do not observe negatively correlated futures prices and the reduction in algebraic complexity of the result is relatively significant.

proof from Watanabe [2008] with the linear multivariate rational expectations form presented in Admati [1985] and Admati and Pfleiderer [1988, 1989].

The interested reader is directed to Watanabe [2008] for a fuller description of the simpler equilibrium conditions as I concentrate on the adjustments needed with the inclusion of segmented trading restrictions. The initial part of this proof illustrates how my model is essentially the same as that contained in Watanabe [2008], page 34-39, which is in turn an adaptation of Admati and Pfleiderer [1989]. I specifically concentrate on the fact that my addition does not lead to the loss of tractability to the analytic solution. Fortunately, the original linear equilibrium for the situation when $\tilde{N}$ is a scalar holds under this restrictive assumption, while it would not hold for the case when $\tilde{N}$ is a positive definite (PD) integer matrix. For this application the restriction is useful and defensible, however future work will focus on this treatment. First, I set up the general form of the asset price progression

$$f(t) = \tilde{\Lambda}_0 + \tilde{\Lambda}_1 \tilde{e} \tag{3.7}$$

following the notation scheme from §(3.2.1) the indexed demand side scheme is

$$\tilde{a}_n = \tilde{B}_0 + \tilde{B}_1 \tilde{\xi}_n \tag{3.8}$$

where the matrices $\tilde{\Lambda}_1 \tilde{e}$ and $\tilde{B}_1$ and the vectors $\tilde{\Lambda}_0 \tilde{e}$ and $\tilde{B}_0$ are computed from the model structure. Using the first filtration

$$\begin{aligned} \tilde{\xi}_n &\equiv \tilde{\mathbb{E}}\left[\hat{\delta} \,|\mathcal{F}_n\right] && (3.9) \\ &= \operatorname{cov}\left(\tilde{\delta}, \tilde{b}_n'\right) \operatorname{var}_t^{-1}\left(\tilde{b}_n\right) \tilde{b}_n && (3.10) \\ &= \tilde{\Sigma}\left(\tilde{\Sigma} + \tilde{\Gamma} + \tilde{\Phi}\right)^{-1} \tilde{b}_n && (3.11) \end{aligned}$$

the profit maximisation condition for each trader is

$$\max_{\tilde{a}_n} \tilde{\mathbb{E}} \left[\left(\tilde{\delta}_T - \tilde{\Lambda}_0 - \tilde{\Lambda}_1 \left(\tilde{a}_n + \sum_{i \neq n} \tilde{a}_i + \tilde{d} \right) \right)' \tilde{a}_n \left| \mathcal{F}_n \right. \right] \tilde{d} \tag{3.12}$$

with system consistent net order flow

$$\tilde{b} = \tilde{a}_n + \sum_{i \neq n} \tilde{a}_i + \tilde{d} \tag{3.13}$$

the asymmetric information in trading process yields a first order condition

$$0 = \tilde{\delta} + \tilde{\xi}_n - \tilde{\Lambda}_0 - \tilde{\Lambda}_1 \left(2\tilde{a}_n + \sum_{i \neq n} \left[\tilde{a}_i \left| \mathcal{F}_n \right. \right] \right) \tag{3.14}$$

consider that this assumption only works if the market is truly partitioned, i.e. an informed trader is restricted to the A market; and this will not hold if the traders have free access to alternate assets within the market. This impacts the second order constraint too, which is now

$$\tilde{\mathbb{E}} \left[\tilde{a}_i \left| \mathcal{F}_n \right. \right] = \tilde{B}_0 + \tilde{B}_1 \tilde{\mathbb{E}} \left[\tilde{\xi}_i \left| \mathcal{F}_n \right. \right] \tag{3.15}$$

where $\tilde{\mathbb{E}} \left[\tilde{\xi}_j \left| \mathcal{F}_n \right. \right]$ is the expectation of another trader $j \in N_i$ estimate of δ. The conditional variance and covariance of the system in 3.11, for any $j \notin n$ is therefore

$$\tilde{\Sigma}_{\tilde{\xi}} \equiv \operatorname{var} \left(\tilde{\xi}_n \right) = \tilde{\Sigma} \left(\tilde{\Sigma} + \tilde{\Gamma} + \tilde{\Phi} \right)^{-1} \tilde{\Sigma} \tag{3.16}$$

$$\tilde{\Sigma}_c \equiv \operatorname{cov} \left(\tilde{\xi}_i, \tilde{\xi}_n' \right) \tag{3.17}$$

$$= \tilde{\Sigma} \left(\tilde{\Sigma} + \tilde{\Gamma} + \tilde{\Phi} \right)^{-1} \left(\tilde{\Sigma} + \tilde{\Gamma} \right) \left(\tilde{\Sigma} + \tilde{\Gamma} + \tilde{\Phi} \right)^{-1} \tilde{\Sigma} \tag{3.18}$$

Then, setting the prior mean $\tilde{\mathbb{E}}(\tilde{\xi}_i) = 0$, the conditional update simplifies to

$$\tilde{\mathbb{E}}\left[\tilde{\xi}_i \,|\mathcal{F}_n\right] = \tilde{\Sigma}_c \tilde{\Sigma}_{\tilde{\xi}}^{-1} \tilde{\xi}_n \tag{3.19}$$

so far this is a repeat of Watanabe [2008], which builds on Admati and Pfleiderer [1988]. However, I now diverge somewhat, as the system consistent first order condition with informed traders will be

$$0 = \tilde{\delta} + \tilde{\xi}_n - \tilde{\Lambda}_0 - 2\tilde{\Lambda}_1 \left(\tilde{B}_0 + \tilde{B}_1 \tilde{\xi}_n\right) - \left(\tilde{N} - 1\right) \tilde{\Lambda}_1 \left[\tilde{B}_0 + \tilde{B}_1 \tilde{\Sigma}_c \tilde{\Sigma}_{\tilde{\xi}}^{-1} \tilde{\xi}_n\right] \tag{3.20}$$

note that the inclusion of $\left(\tilde{N} - 1\right) \tilde{\Lambda}_1 \left[\tilde{B}_0 + \tilde{B}_1 \tilde{\Sigma}_c \tilde{\Sigma}_{\tilde{\xi}}^{-1} \tilde{\xi}_n\right]$, holds only if $tr\tilde{N} = \tilde{N}$ under the Watanabe [2008] approach. This now solves for the system of equations in 3.4 to 3.6, first set

$$0 = \tilde{I} - 2\tilde{\Lambda}_1 \tilde{B}_1 - \left(\tilde{N} - 1\right) \tilde{\Lambda}_1 \tilde{B}_1 \tilde{\Sigma}_c \tilde{\Sigma}_{\tilde{\xi}}^{-1}, \quad or \tag{3.21}$$

$$\tilde{\Lambda}_1 \tilde{B}_1 = \tilde{J}^{-1}, \quad with \quad \tilde{J} \equiv 2\tilde{I} + \left(\tilde{N} - 1\right) \tilde{\Sigma}_c \tilde{\Sigma}_{\tilde{\xi}}^{-1} \tag{3.22}$$

back substitution from 3.16 and 3.17, reproduces the expression suggested for J in 3.6. Again I see the Hadamard product enter the system,

$$\tilde{\Lambda}_0 = \tilde{\delta} - \left(\tilde{N} - 1\right) \tilde{\Lambda}_1 \tilde{B}_0 \tag{3.23}$$

following from this, the market maker efficiency condition will be

$$f(t) = \tilde{\delta} + \tilde{\mathbb{E}}\left[\tilde{\delta} \,|\mathcal{F}\theta_m\right] = \tilde{\delta} + \text{cov}\left(\tilde{\delta}, \tilde{e}'\right) \text{var}^{-1}\left(\tilde{e}\right) \left(\tilde{e} - \tilde{N}\tilde{B}_0\right) \tag{3.24}$$

proceeding in the standard fashion reveals the market maker variance covariance

expression

$$\text{cov}\left(\tilde{\delta},\tilde{e}'\right)=\sum_{i=1}^{\tilde{N}_i}\text{cov}\left(\tilde{\delta},\tilde{\delta}_i'\right)\tilde{B}_1'=\tilde{N}\tilde{\Sigma}_{\tilde{\xi}}\tilde{B}_1' \tag{3.25}$$

by extension

$$\begin{aligned}\text{var}\left(\tilde{e}\right) &= \tilde{B}_1\text{var}\left(\sum_{i=1}^{\tilde{N}_i}\tilde{\xi}_n\right)\tilde{B}_1'+\tilde{\Psi}\\ &= \tilde{N}\tilde{B}_1\left\{\tilde{\Sigma}_{\tilde{\xi}}+\left(\tilde{N}-1\right)\tilde{\Sigma}_{\tilde{c}}\right\}\tilde{B}_1'+\tilde{\Psi}\equiv\tilde{\Sigma}_{\tilde{e},\tilde{c}}\end{aligned} \tag{3.26}$$

once again from standard matrix commutation rules this holds, only under the condition that $tr\tilde{N}=\tilde{N}$, this yields the liquidity adjusted pricing formula to be

$$\begin{aligned}f(t) &= \tilde{\delta}+\tilde{N}\tilde{\Sigma}_{\tilde{\xi}}\tilde{B}_1'\left[\tilde{N}\tilde{B}_1\left\{\tilde{\Sigma}_{\tilde{\xi}}+\left(\tilde{N}-1\right)\tilde{\Sigma}_c\right\}\tilde{B}_1'+\tilde{\Psi}\right]^{-1}\\ &\quad\times\left(\tilde{e}-\tilde{N}\tilde{B}_0\right) \qquad (3.27)\\ &= \tilde{\delta}+\left[\tilde{B}_1\left\{\tilde{I}+\left(\tilde{N}-1\right)\tilde{\Sigma}_c\tilde{\Sigma}_{\tilde{\xi}}^{-1}\right\}+\tilde{N}^{-1}\tilde{\Psi}\tilde{B}_1'^{-1}\tilde{\Sigma}_{\tilde{\xi}}^{-1}\right]^{-1}\\ &\quad\times\left(\tilde{e}-\tilde{N}\tilde{B}_0\right) \qquad (3.28)\end{aligned}$$

again I proceed along the lines of Watanabe [2008], generalising from Admati and Pfleiderer [1988] with the addition of my separated trader condition,

$$\tilde{\Lambda}_1 = \left[\tilde{B}_1\left\{\tilde{I}+\left(\tilde{N}-1\right)\tilde{\Sigma}_c\tilde{\Sigma}_{\tilde{\xi}}^{-1}\right\}+\tilde{N}^{-1}\tilde{\Psi}\tilde{B}_1'^{-1}\tilde{\Sigma}_{\tilde{\xi}}^{-1}\right]^{-1},\quad or \tag{3.29}$$

$$\tilde{\Lambda}_1\tilde{B}_1 = \left[\tilde{J}-\tilde{I}+\tilde{N}^{-1}\tilde{B}_1^{-1}\tilde{\Psi}\tilde{B}_1'^{-1}\tilde{\Sigma}_{\tilde{\xi}}^{-1}\right]^{-1},\quad and \tag{3.30}$$

$$\tilde{\Lambda}_0 = \tilde{\delta}-\tilde{N}\tilde{\Lambda}_1\tilde{B}_0 \tag{3.31}$$

eliminating $\left(\tilde{N}-1\right)\tilde{\Sigma}_c\tilde{\Sigma}_{\tilde{\xi}}^{-1}$ leaves a system of four equations for the four unknowns $\tilde{\Lambda}_1\tilde{e}$ and $\tilde{B}_1$ and the vectors $\tilde{\Lambda}_0\tilde{e}$. Cancelling for $\tilde{J}$ yields

$$\tilde{B}_1^{-1}\tilde{\Psi}\tilde{B}_1'^{-1}=\tilde{N}\tilde{\Sigma}_{\tilde{\xi}} \tag{3.32}$$

substituting for my standard functional form of $\tilde{B}$, whereby $\tilde{B}_1 = \tilde{\Lambda}_1^{-1}\tilde{J}^{-1}$, I find

$$\tilde{\Lambda}_1\tilde{\Psi}\tilde{\Lambda}_1 = \tilde{N}\tilde{J}^{-1}\tilde{\Sigma}_{\tilde{\xi}}\tilde{J'}^{-1} \tag{3.33}$$

as such

$$\left(\tilde{\Psi}^{\frac{1}{2}}\tilde{\Lambda}_1\tilde{\Psi}^{\frac{1}{2}}\right)^2 = \tilde{N}\tilde{\Psi}^{\frac{1}{2}}\tilde{J}^{-1}\tilde{\Sigma}_{\tilde{\xi}}\tilde{J'}^{-1}\tilde{\Psi}^{\frac{1}{2}} \equiv \tilde{N}\tilde{M} \tag{3.34}$$

element by element dividing by $\tilde{N}$ yields an expression for $\tilde{M}$, which is of the form VUV' and therefore PD. Therefore a Cholesky factor $\tilde{M}^{\frac{1}{2}}$ exists. The resultant order flow is therefore

$$\tilde{a}_n = \tilde{B}_1\tilde{\xi}_n = \tilde{\Lambda}_1^{-1}\tilde{J}^{-1}\tilde{\Sigma}\Big(\tilde{\Sigma} + \tilde{\Gamma} + \tilde{\Phi}\Big)^{-1}\tilde{b}_n = \tilde{\Lambda}_1^{-1}\tilde{J}^{-1}\tilde{\Sigma}_{\tilde{\xi}}\tilde{\Sigma}^{-1}\tilde{b}_n \tag{3.35}$$

The matrix $\tilde{\Lambda}_1^{-1}\tilde{J}^{-1}\tilde{\Sigma}\Big(\tilde{\Sigma} + \tilde{\Gamma} + \tilde{\Phi}\Big)^{-1}$ is the solution to the autoregressive terms of B_1, given the structure of the solution to $\tilde{\Lambda}_1 = \tilde{\Lambda}$, which is PD then we can see that the multivariate equivalence of Kyles lambda is a PD matrix autoregressive coefficient. Hence the mapping of time-varying vector autoregressive model (TV-VAR), whereby the Grammian matrix $(X'X)^{-1}$ should be partitioned as a function of $\tilde{\Lambda}$. This is very useful as it allows me to show that the liquidity augmented price is linear in liquidity shocks.■

I now report two market equilibrium artefacts as propositions that rearrange the equilibrium equations of state, the expected quadratic variation and the expected volume of trade across the two assets. It is useful to use these to guide the calibration of the model.

Proposition 1. *Expected Price Variance Let* $\Delta t\tilde{H} = \mathbb{E}[(f(t+\Delta t) - f(t))(f(t+\Delta t) - f(t))']$ *be the observed pricing variance-covariance matrix. From Theorem*

1, this is determined from $\tilde{\Theta} \in \{\tilde{\Sigma}, \tilde{\Gamma}, \tilde{\Phi}, N_B\tilde{\Psi}\}$ *by the following expression*

$$\Delta t\tilde{H} = \tilde{\Sigma} + \tilde{\Sigma}_{\Lambda}, where \quad \tilde{\Sigma}_{\Lambda} = N_A\tilde{\Sigma}((N_A+1)(\tilde{\Sigma}+\tilde{\Gamma}) + 2\tilde{\Phi})^{-1}\tilde{\Sigma} \tag{3.36}$$

From Theorem 1, the variance of the price process is mechanistically specified from the multivariate autoregression as

$$\Delta t\tilde{H} \equiv \text{var}\left(f(t+\Delta t) - f(t)\right)) = \text{var}\left(\tilde{\Lambda}\tilde{e}\right) + \text{var}\left(\tilde{\delta} - \tilde{\Lambda}\tilde{e}\right) \tag{3.37}$$

by induction the variance follows from the noise of the submission process and the multivariate extension of Kyle's lambda, therefore the variance iteration will be

$$\text{var}\left(\tilde{\Lambda}\tilde{e}\right) = \tilde{\mathbb{E}}\left[\tilde{\Lambda}\tilde{e}\tilde{e}'\tilde{\Lambda}\right] = \tilde{\mathbb{E}}\left[\tilde{\Lambda}\tilde{\mathbb{E}}\left(\tilde{e}\tilde{e}'\right)\tilde{\Lambda}\right] = \tilde{\mathbb{E}}\left[\tilde{\Lambda}\text{var}\left(\tilde{e}\right)\tilde{\Lambda}\right] \tag{3.38}$$

Simple rearrangement and substitution from the definitions in Theorem 1 yields

$$\text{var}\left(\tilde{e}\right) = \tilde{\Lambda}^{-1}\text{cov}\left(\tilde{\delta}, \tilde{e}'\right) = \tilde{N}\tilde{\Lambda}^{-1}\tilde{\Sigma}_{\tilde{\xi}}\tilde{B}_1' = \tilde{N}\tilde{\Lambda}^{-1}\tilde{\Sigma}_{\tilde{\xi}}\tilde{J'}^{-1}\tilde{\Lambda}^{-1} \tag{3.39}$$

and therefore the expectation collapses to

$$\text{var}\left(\tilde{\Lambda}\tilde{e}\right) = \tilde{\mathbb{E}}\left[\tilde{N}\tilde{\Sigma}_{\tilde{\xi}}\tilde{J'}^{-1}\right] \tag{3.40}$$

I show that the imposition of $\tilde{N} = [N_i]$ runs through the derivation without loss of generality, therefore

$$\tilde{N}\tilde{\Sigma}_{\tilde{\xi}}\tilde{J'}^{-1} = \tilde{N}\left(\tilde{J'}\tilde{\Sigma}_{\tilde{\xi}}^{-1}\right)^{-1} = \tilde{N}\left[2\tilde{\Sigma}_{\tilde{\xi}}^{-1} + \left(\tilde{N}-1\right)\tilde{\Sigma}_{\tilde{\xi}}^{-1}\tilde{\Sigma}_{\tilde{c}}\tilde{\Sigma}_{\tilde{\xi}}^{-1}\right]^{-1} \tag{3.41}$$

The diagonality condition imposed on $\tilde{N}$ reduces the noise to the following diag-

onal matrix

$$\tilde{N}\tilde{\Sigma}_{\tilde{\xi}}\tilde{J'}^{-1} = \tilde{N}\left[2\tilde{\Sigma}_{\tilde{\xi}}^{-1} + \left(\tilde{N}-1\right)\tilde{\Sigma}^{-1}\left(\tilde{\Sigma}+\tilde{\Gamma}\right)\tilde{\Sigma}^{-1}\right]^{-1} \tag{3.42}$$

Substitution from Theorem 1 definitions yields

$$\tilde{N}\tilde{\Sigma}_{\tilde{\xi}}\tilde{J'}^{-1} = \tilde{N}\tilde{\Sigma}\left[\left(\tilde{N}+1\right)\left(\tilde{\Sigma}+\tilde{\Gamma}\right)+2\tilde{\Phi}\right]^{-1}\tilde{\Sigma} \equiv \tilde{\Sigma}_{\tilde{\Lambda}\tilde{e}} \tag{3.43}$$

Add up all of the variances and impose diagonal restrictions on the covariances to derive the variance condition,

$$\begin{aligned} \text{var}\left(\tilde{\delta}-\tilde{\Lambda}\tilde{e}\right) &= \tilde{\Sigma} - \text{cov}\left(\tilde{\delta},\tilde{e}'\right)\tilde{\Lambda} - \left[\text{cov}\left(\tilde{\delta},\tilde{e}'\right)\tilde{\Lambda}\right]' && (3.44)\\ &\quad +\text{var}\left(\tilde{\Lambda}\tilde{e}\right) \\ &= \tilde{\Sigma} - \tilde{\Sigma}_{\tilde{\Lambda}\tilde{e}} && (3.45) \end{aligned}$$

therefore yielding the volatility expectation. ■

Proposition 2. *Expected Market Volume Following Admati [1985], Admati and Pfleiderer [1988] and Admati and Pfleiderer [1989] let aggregate volume be defined by*

$$\tilde{V} = \tfrac{1}{2}|\sum_{n=1}^{N_A}\tilde{a}_n| + |\tilde{d}| + |\tilde{e}| \tag{3.46}$$

in terms of the equilibrium in Theorem 1 this is given by:

$$\begin{aligned} \mathbb{E}_0[\tilde{V}] &= \tfrac{1}{\sqrt{2\pi}}\mathbb{E}_0[(diag(N_A\tilde{B}(N_A(\tilde{\Sigma}+\tilde{\Gamma})+\tilde{\Phi})\tilde{B})^{1/2} + && (3.47)\\ &\quad diag(\Psi))^{1/2} + (diag(N_A\tilde{B}(N_A(\tilde{\Sigma}+\tilde{\Gamma})+\tilde{\Phi})\tilde{B}' + N_B\tilde{\Psi}))^{1/2}] \end{aligned}$$

I utilize Lemma 2 from Watanabe [2008] to begin with augmenting it by replacing

$\left(\tilde{B}\sum_{n=1}^{\tilde{N}_i}\tilde{b}_n\right)$ with $\left(\tilde{B}\sum_{n=1}^{tr\tilde{N}_i}\tilde{b}_n\right)$. From this the net order flow in my notation will be,

$$\begin{aligned}\tilde{\Sigma}_{\tilde{e}} &= \text{var}\left(\tilde{e}\right) = \text{var}\left(\tilde{B}\sum_{n=1}^{tr\tilde{N}_i}\tilde{b}_n\right) + \text{var}\left(\tilde{d}\right) \\ &= \tilde{N}\tilde{B}\left[\tilde{N}\left(\tilde{\Sigma}+\tilde{\Gamma}\right)+\tilde{\Phi}\right]\tilde{B}'+\tilde{\Psi} \end{aligned} \tag{3.48}$$

therefore in expectations the volume will be

$$\tilde{\mathbb{E}}_0\left[\tilde{V}\right] = \frac{1}{2}\tilde{\mathbb{E}}_0\left[\tilde{\mathbb{E}}\left|\sum_{n=1}^{tr\tilde{N}_i}\tilde{a}_n\right| + \tilde{\mathbb{E}}\left|\tilde{d}\right| + \tilde{\mathbb{E}}\left|\tilde{e}\right|\right] \tag{3.49}$$

The last part is identical to Lemma 2 of Watanabe [2008], except with the inclusion of the Hadamard product expression, under my diagonal specification, it is easy to see that this is in effect two separate Watanabe [2008] models. This is sequentially simplified to

$$\begin{aligned}\text{var}\left(\tilde{a}_n\right) &= \tilde{B}\left(\tilde{\Sigma}+\tilde{\Gamma}+\tilde{\Phi}\right)\tilde{B}' && (3.50)\\ &= \tilde{\Lambda}^{-1}\tilde{J}^{-1}\tilde{\Sigma}_{\tilde{\xi}}\tilde{\Sigma}^{-1}\left(\tilde{\Sigma}+\tilde{\Gamma}+\tilde{\Phi}\right)\tilde{\Sigma}^{-1}\tilde{\Sigma}_{\tilde{\xi}}^{-1}\tilde{\Sigma}_{\tilde{\xi}}\tilde{J'}^{-1}\tilde{\Lambda}^{-1} && (3.51)\\ &= \tilde{\Lambda}^{-1}\tilde{J}^{-1}\tilde{\Sigma}_{\tilde{\xi}}\tilde{J'}^{-1}\tilde{\Lambda}^{-1} = \tilde{N}^{-1}\tilde{\Psi} && (3.52)\end{aligned}$$

recalling the definitions from Proposition 1 and 2 and hence the cross market volume. ■

3.2.2 Existence of a Non-linear Response in the Price Impact Matrix $\tilde{\Lambda}$

On average I observe between 100 and 200 traders in the market at any given moment. The term structure of trading activity tends to be at a plateau for contracts between 0 and 3 years from maturity and drops towards single digits past 5 years. Following data from the CME group, I set N_B to be 20 and adjust the number of N_A traders from zeros to the majority, keeping N_B fixed. For the parameters, I use the following as a baseline specification for my simulation. I do not have specific point estimates for $\tilde{\Gamma}$, $\tilde{\Psi}$ and $\tilde{\Phi}$, however, if I set the first element of $\tilde{\Sigma}$ to be the square of the approximate long run daily volatility of the reference LIBOR rate and $\tilde{H}$ and $\tilde{V}$ to be the observed price volatility and volume of the future I can use these reference to points to generate reasonable domains for $\tilde{\Gamma}$, $\tilde{\Psi}$ and $\tilde{\Phi}$.

$\tilde{\Sigma}$, the variance covariance matrix of the true valuation price process $\sigma_1 = 1/4$, $\sigma_3 = \sigma_1/5$ and $\sigma_2 = 7/10\sigma_1\sigma_3$.

$\tilde{\Gamma}$, the covariance matrix of the global noise across all A traders signals $\gamma_1 = 1/10$, $\gamma_3 = 1/10$ and $\gamma_2 = 5/10\gamma_1\gamma_3$.

$\tilde{\Psi}$ the covariance matrix for each A traders signal $\psi_1 = \{1/4, 1/2, 1, 3/2\}$, $\psi_3 = \psi_1/2$ and $\psi_2 = 0.2\psi_1\psi_3$.

$\tilde{\Phi}$ the variance of an individual B traders order flow $\phi_1 = 1$, $\phi_3 = \phi_1/2$ and $\phi_2 = 0$.

In Figure 3.1, I plot the level ($diag\tilde{\Lambda}$) and the partial derivative $\partial\ diag\tilde{\Lambda}/\partial n_A$ of the elements price impact matrix $\tilde{\Lambda}$ that determines the degree of misalignment

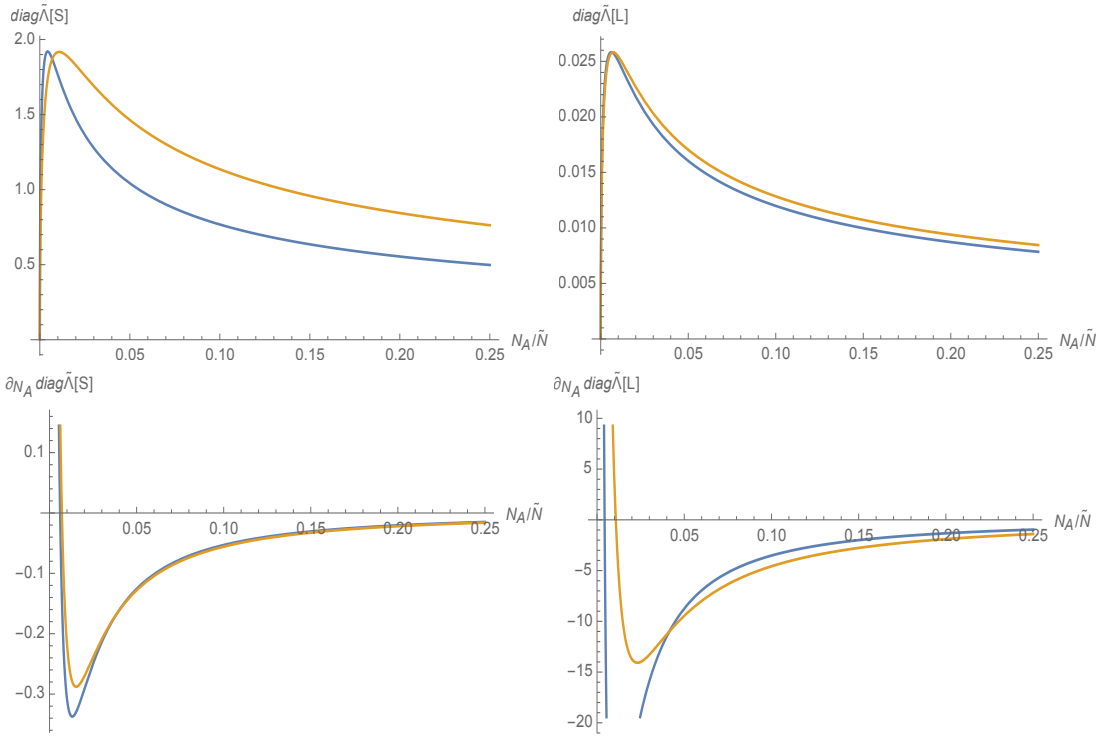


Figure 3.1: The Effect in Level and Derivative of a Change in the Fraction of Type A Traders versus Type B Traders

Notes: I do not model explicitly the advantage that speed gives the traders other than I presume that the ability to rapidly anticipate the direction of the underlying value process from the order flow provides a systematic advantage. Hence, Type A has a noisy but unbiased expectation of the terminal valuation $\tilde{\delta}(T)$. The model postulates four sources of quadratic variation in the market. $\tilde{\Sigma}$ is the quadratic variation in the underlying asset. $\tilde{\Gamma}$ is the global noise disturbing all type Type A traders forward looking signals. $\tilde{\Phi}$ represents the quadratic variation of the idiosyncratic noise disturbing each of the N_A Type A traders signal over and above the global noise. Finally, $\tilde{\Phi}$ is the variance-covariance matrix describing the quadratic variation of each of the N_B traders random submissions. In the four quadrants, I plot the diagonal elements of the resulting $\tilde{\Lambda}$ matrix from the market clearing equilibrium in Theorem 1 and its derivative with respect to $N_A/\tilde{N}$ where $\tilde{N} = N_A + N_B$. Recalling that $f(T) = \delta(T) + \tilde{\Lambda}\tilde{e}$, where $\tilde{e}$ is the aggregate net order flow.

between the observed futures price $f(T)$ and $\delta(T)$ as a function of aggregate order flow $\tilde{e}$ with respect to the fraction $N_A/\tilde{N} \in [0.1, 0.8]$ of Type A traders in the market. We can think of the elements of the $\tilde{\Lambda}$ as the representative of the depth of the market and the impact of order-flow on price.

We can see that across a variety of values of ψ_1 and ψ_2 that the model predicts that a small number of Type A traders will increase the magnitude of the first S and second L elements of $diag\tilde{\Lambda}$, however, as the number of Type A traders increases the detrimental effect on the market decreases and rapidly reverses. I will now introduce some previous studies and then move on to demonstrating that the pattern of market depth as a function of the fraction of traders advantaged by speed increases has a similar pattern to that predicted herein.

3.3 A Brief Review of High Frequency Trading and Market Liquidity

The prior literature investigated that algorithmic trading can increase liquidity and market efficiency; hence, it can accelerate to the price discovery function. Others suggested that AT has effects on increasing price volatility and reducing liquidity, especially when financial markets are in times of stress.

Frino and McKenzie [2002] put forward that the screen trading can enhance the price discovery in both spot and future markets. Besides, they also discuss that the futures lead the spot market in price discovery, this has been supported by Harris [1989], Stoll and Whaley [1990], Chan [1992] and Huang and Stoll [1994]. Engle et al. [2012] directly adopt the detailed order-execution data from Morgan Stanley, which includes the arrival price, parent trades size, and volume-weighted

average price. They employ these data to measure the impacts on execution risk and cost of algorithm aggressiveness. Domowitz and Yegerman [2006] use ITG's transaction cost for over 40 buy-side institutions to examine performance from different algorithmic trading services. They believe that algorithmic trading is a cost-effective technique, however, their empirical evidence "suggests that algorithms are not yet sophisticated enough for large order sizes" [Domowitz and Yegerman, 2006, p. 11].

Despite the equity markets, AT platform has been widely spread to futures, foreign exchange and commodities markets. CME Group [2010] investigates the effects of algorithmic trading methodologies on market liquidity in a series of CME products, such as Eurodollar, EuroFX, E-mini S&P 500 and Crude Oil futures. CME Group [2010] asks their clients to register their Operator ID (Tag 50) through Automated Trading Systems (ATS) in 2006, therefore, they can monitor the volume of AT orders relative to the total trades volume, even the volume of electronic message traffic. They report that the volume and message traffic originating from the increased AT activities is associated with increased liquidity and reduced volatility. While, the results according to different markets are not balanced. The results are relatively positive for Eurodollar, Crude Oil future and 10-Year Treasury note markets. E-mini S&P 500 market shows mixed results. As for the foreign exchange markets, Chaboud et al. [2014] analyze the relation between AT and price discovery process with a more comprehensive dataset. They employ Euro-Dollar, Dollar-Yen and Euro-Yen trading data from Electronic Broking Services (EBS). Their data set has been clearly classified whether the transactions belongs to human or computer trades. They put forward that algorithmic traders are more following highly correlated strategies. Furthermore, there's little evidence about the relation between AT and volatility, which is consistent with Hendershott et al. [2011]. However, Chaboud et al. [2014] illustrate

that most volatility in FX returns should attribute to non-algorithmic orders.

There is a wide range of studies on the effects of market liquidity on the price information process [Harris, 1990; Kyle, 1985; Madhavan, 2000; O'Hara, 1995]. Kyle [1985] points out that three key characteristics of liquidity can embody the market microstructure: Tightness, Depth, and Resilience. Tightness can be measured by the gaps of bids and asks; Depth can be measured by the trades large quantities without changing current price; Resilience refers the ability to trade assets with little influence on the price. My study considers to use bid-ask spreads, quoted half-spread, effective half-spread, realized spread, price impact and quoted depth as liquidity indicators.

As an important component of transaction costs, the bid-ask spread is a commonly used indicator of market liquidity. A smaller bid-ask spread indicates a more liquid market; otherwise, a larger bid-ask spread implies poor liquidity. There are many ways to measure bid-ask spreads, including quoted spreads, effective spreads and realized spreads. Quoted spread is simply the difference between the bid prices and ask prices. Effective spread can be calculated by the trade prices comparing the midpoint of bid and ask at the time of the execution. Considering the impacts of transaction on the prices, effective spread can indicate the deviation between the actual price and the mid price before the execution. Price impact is calculated as the mid-quote of the bid and offer at trade time comparing with the mid-quote at a some time period after trade time. Compared with effective spread, realized spread is the difference between trades prices and the midpoint of bids and asks at a specified time period after the time of the execution. The existence of informed traders will lead to market prices to shift up after buying or down after selling. The realized spread considers and captures these adverse price movements. The relation among these spreads is that effective

spread can be equal to the sum of price impact and realized spread.

Since HFT has increased sharply over the past decade, some studies have demonstrated direct evidence that the automated trading results in a narrower bid-ask spread [Angel et al., 2011; Bershova and Rakhlin, 2013; Carrion, 2013; Castura et al., 2010; Hendershott et al., 2011; Jovanovic and Menkveld, 2010; Malinova et al., 2013; Menkveld and Zoican, 2014; Zhang and Riordan, 2011]. Hendershott et al. [2011] investigate the influence of the increasing algorithmic trading from 2001 to 2005 on different measures of spreads. They find a significant reduction in effective spreads across all market-cap quintiles, with most quintiles decreasing to 1/3 of their start levels. Angel et al. [2011] examine the measurements of market quality and how they change over time and in response to regulatory changes and structural changes in the US equity markets. They state that both execution speeds and trading costs have fallen, bid-ask spreads have narrowed, and available liquidity has increased.

A similar conclusion has also been put forward by Castura et al. [2010] who, examining the US equities market find that the bid-ask spread has narrowed, available liquidity has increased and price efficiency has improved. Menkveld [2013] provides direct evidence of a link between a single, large, high-frequency market maker and the narrowing of spreads on the Chi-X and Euronext from January 2007 to June 2008. Over this period, bid-ask spreads in Dutch stocks declined by 50% relative to Belgian stocks, which implies that high-frequency market making reduces bid-ask spreads. The relation between HFT and different spreads has been discovered to depend on whether HFT is categorized as an aggressive activity or a passive activity. HFT tends to trade passively when spreads are wide and trade aggressively when spreads are narrow [Carrion, 2013; Zhang and Riordan, 2011]. Castura et al. [2010] also put forward that realized

spreads are significantly more negative when HFTs trade aggressively.

In addition to the bid-ask spreads, market depth is an important indicator of liquidity. Recent studies have reported evidence to suggest strong, positive impacts of HFT on the market liquidity [see Angel et al., 2011; Brogaard, 2010; Castura et al., 2010; Jarnecic and Snape, 2014]. Market liquidity indicates the ability for investors to obtain ideal inventory positions with low execution costs and low transactional price impacts. Angel et al. [2011] estimate that the median depth at the NBBO increased across all the US stocks. Furthermore, when allocating their portfolios to diverse stocks, investors need to convert liquidity into dollar amounts; therefore, the average dollar amount of stock quoted is a better measure of liquidity. Whilst this type of analysis is very common for stock, I can find very little evidence for the prior analysis of inside-spreads for money market/interest rate derivatives. An alternative equity market study, Castura et al. [2010] uses the dollar amount available at the inside-spread to measure the available liquidity at any instant in time and compute quarterly time averages. The results show a reduction in available liquidity during the financial crisis of 2007/2008 and available liquidity that reverted to a high level in 2010. A similar study by Jarnecic and Snape [2014] employ data from the London Stock Exchange (LSE) and find that high-frequency traders tend to provide liquidity at large spreads and to demand liquidity at narrow spreads. Additionally, high-frequency traders are likely to smooth out liquidity over time; however, their impact on volatility is uncertain.

The most closely related to this study are the papers by Hendershott et al. [2011] and Hendershott and Riordan [2013]. Hendershott et al. [2011] investigate the relation between AT and liquidity using NYSE stocks over the five years between February 2001 and December 2005. Using Biais and Weill [2009] for references,

Hendershott et al. [2011] employ the NYSE electronic message traffic as a proxy for AT. The message traffic includes order submissions, cancellations and trade reports. Because of the introduction of 'Autoquote' at NYSE in 2003, they also use this market structure change as an instrument variable to assess the impacts of AT on liquidity. Hendershott et al. [2011] indicate that AT can narrow quoted spread, and reduce price impacts and trade-related price discovery, especially for stocks with large market capitalization. The finding illustrates that AT enhances liquidity and makes quotes more informative.

However, using message traffic as AT proxy is difficult to directly check "when and how AT behave and their role in liquidity supply and demand" [Hendershott and Riordan, 2013, p. 1004]. On the other hand, Hendershott et al. [2011] use Autoquote as an instrumental variable to evaluate that AT has the effects to enhance liquidity and makes quotes more informative. While, Hendershott and Riordan [2013] focus on the detecting the channels by which AT could increase market liquidity.

3.4 Empirical Measurements

My analysis is divided into two parts. First, I generate a series of market quality indicators from high frequency data. In the main, I use the bids and asks in the entire limit order book and transaction data for this information, and I then aggregate this to daily data. Second, I generate daily approximations of the fractions of high frequency algorithmic trading (as the speed gain will be specifically via a computerized algorithm submitting quotes via an electronic communication network).

3.4.1 Liquidity Spreads Measurements

My liquidity measures follow the approach of Hendershott et al. [2011] which provides a comprehensive overview of the various liquidity measures from the literature. They only have the inside quotes and transactions data. Inside quotes only include the highest bid prices and lowest ask prices, it can not provide the complete picture of order flow. My data contains the whole limited order book and trades, so it is possible to calculate liquidity spreads for each level. Moreover, Hendershott et al. [2011] use the Lee and Ready [1991] algorithm to calculate trade direction, however, this method is not that accurate for HFT data. I found out about 85% belongs to "no trades" classification. Therefore, I alter the Lee-Ready Algorithm using the volume weighted average price of five-level order book. In total, I measure 8 different liquidity spreads and use them as dependent variables in regression, bid-ask spreads ($\tilde{\mathscr{S}}^{Q}$), quoted half-spread ($\tilde{\mathscr{S}}^{Q^{1/2}}$), quoted depth ($\tilde{\mathscr{S}}^{D}$), effective half-spread ($\tilde{\mathscr{S}}^{E}$), realized spreads ($\tilde{\mathscr{S}}^{R,5}$ for 5 minutes and $\tilde{\mathscr{S}}^{R,30}$ for 30 minutes) and price impacts ($\tilde{\mathscr{S}}^{AS,5}$ for 5 minutes and $\tilde{\mathscr{S}}^{AS,30}$ for 30 minutes).

This section I follow the timing convention from §(2.1) that t represents the length of trading time on an individual day, k denotes an intraday time index and $j \in \{1, 2, \ldots, 5\}$ is the order book level. Bid-ask spreads and quoted half-spread are simply the difference between the bid prices and ask prices. Quoted half spreads, $\tilde{\mathscr{S}}^{Q^{1/2}}_{jk}$, can be defined as

$$\tilde{\mathscr{S}}^{Q^{1/2}}_{jk} = 100(p_{ajk} - p_{bjk})/2(p_{mjk}) \tag{3.53}$$

where j is the j^{th} level of the order book, p_{ajk} and p_{bjk} are the ask price and bid price for level $j = \{1, 2, ..., 5\}$ at intraday time k, and p_{mjk} is the quoted

mid-price prevailing at the time of the k_{th} trade. Quoted depth indicates the competition and capacity of the market. Larger quoted depth means there are more competitors and the market can absorb more orders before changing the current price. I calculate the five levels of the order book quoted depth, $\tilde{\mathscr{S}}^D_{jk}$, to have greater accuracy

$$\tilde{\mathscr{S}}^D_{jk} = p_{ajk}v_{ajk} + p_{bjk}v_{bjk} \tag{3.54}$$

where v_{ajk} and v_{bjk} are the ask volume and bid volume for level j at intraday time k.

The effective half-spread shows the difference between the actual trades prices and the mid-quote prices. The smaller the effective spreads, the more liquid is the contract and the more actual transaction prices are near the mid prices. For the k_{th} trade relative to the quote level j, the proportional effective half-spread, $\tilde{\mathscr{S}}^E_{jk}$, can be expressed as

$$\tilde{\mathscr{S}}^E_{jk} = q_{jk}(p_{jk} - p_{mjk})/p_{mjk}, \tag{3.55}$$

where q_{jk} is trade direction indicator that +1 for buyer-initiated trades, 0 for no trades, and -1 for seller-initiated trades; p_{jk} is the actual transaction price at time k; and m_{jk} is the quoted mid-price at the time of the k_{th} trade. Unlike effective spreads, realized spreads describe the difference between quoted mid-price and the actual transaction at a specified time period after the trade execution time. Hence, the realized spread can capture the adverse price shifts caused by the existence of informed traders. Here I calculate realized spread at 5 minutes and 30 minutes after the trades. The realized spread (5 mins) for the k_{th} transaction at the j quote level is defined as

$$\tilde{\mathscr{S}}^{R,5}_{jk} = q_{jk}(p_{jk} - p_{m,j,k+5min})/p_{mjk}, \tag{3.56}$$

where p_{jk} is the actual transaction price at the time of the k_{th} trade, q_{jk} is the buy-sell trade direction indicator, m_{jk} is the midpoint prevailing at the time of the k_{th} trade, and $m_{j,k+5min}$ is the quoted mid-price at 5 minutes after the k_{th} trade. Similarly, the 30-minute realized spread can be measured analogously with the quoted mid-price at 30 minutes after the transaction ($m_{j,k+30min}$).

$$\tilde{\mathscr{S}}_{jk}^{R,30} = q_{jk}(p_{jk} - p_{m,j,k+30min})/p_{mjk}, \tag{3.57}$$

To evaluate the impact of informed traders shifting market prices, I calculate the 5 minutes and 30 minutes price impact of a trade. The adverse selection, $\tilde{\mathscr{S}}_{jk}^{AS,5}$, is calculated as the difference between the quoted mid-price of the bid and offer at trade time and the quoted mid-price at 5 minutes after execution,

$$\tilde{\mathscr{S}}_{jk}^{AS,5} = q_{jk}(p_{m,j,k+5min} - p_{mjk})/p_{mjk}, \tag{3.58}$$

Similarly, the 30-minute adverse selection is defined as

$$\tilde{\mathscr{S}}_{jk}^{AS,30} = q_{jk}(p_{m,j,k+30min} - p_{mjk})/p_{mjk}, \tag{3.59}$$

The arithmetic identity among these spreads is that effective spread can be equal to the sum of price impact (adverse selection) and realized spread.

$$\tilde{\mathscr{S}}_{jk}^{E} = \tilde{\mathscr{S}}_{jk}^{R} + \tilde{\mathscr{S}}_{jk}^{AS} \tag{3.60}$$

3.4.2 A Proxy For the Fraction of High Frequency Algorithmic Trading in the Market

The objective of this chapter is to understand the impact of speed on the preceding market quality indicators. My dataset consists of the entire limit order book from inside quotes to the highest quote level (usually level 5 for the CME), sorted by price. I believe that this is the first large scale study of a complete limit order book ever attempted. One of the useful aspects of analyzing Eurodollar futures is that the limit order book and trades from the Globex trading platform provides the entire picture of activity in the market. Eurodollars are not cleared through CMEs over-the-counter (OTC) trading platform (like light crude or currency futures) and since the closure of open outcry trading pits the Globex platform process all trades and quotes.

Whilst I have all of the available message updates for the Eurodollar market the data does not show directly whether orders come from computer algorithm or human being, that is I do not have the trading account numbers (each has a prefix that determines if the account is an AT and hence generates HFT order flow). However, I have a couple of guides as to the approximate average fraction of ATs in the market. First, the CME groups own study, CME Group [2010], with access to the account numbers for a snapshot of trading in 2010 reports that around 64% of Eurodollar messaging activity on the limit order book is generated by ATs. Second, I have access to the number of accounts actually active on the exchange. For instance, if ten accounts are active and a human being can generate new orders at speeds of approximately 500ms then the likelihood of two non AT generated orders arriving within 25ms of each other is relatively low. Therefore HFT as a pattern of trades that arrives on the order book within a very short time period of each other (conditional on the number of traders in the market)

should provide a proxy to the quantity of order flow ascribed to HFT and hence the number of ATs in the market.

The physiology literature estimates the average human reaction time as being, in general, greater than 250ms [Kosinski, 2008]. However, it is unlikely that a human being over the whole trading day will generate a new quote on a 250ms grid. For my main AT proxy, I threshold at 25ms as this yields a modal fraction of order flow messages close to the average quantity of order flow computed by CME with access to the AT account prefixes, see Figure 3.4. For robustness, I have also constructed proxies at between 50ms and 200ms and the time evolution of these proxies closely correlates to the 25ms proxy (on the most active days 90% of messages are below 200ms, however, 75% are also below 25ms indicating that the clustering is no accident of numbers of traders). Furthermore, I have the number of traders actively quoting in the market at any given time (the number of active connections to the exchange) on the very busiest days no more than 200 participants are quoting in any given contract and for the majority of the time this is substantially less than 50. I precondition, non-parametrically, the total active number of traders in the market as an instrumental variable to extract the endogeneity that might be inherent in construction of the proxy.

Hence, let the algorithmic trading messages, $\tilde{\mathcal{M}}^{\tilde{\mathcal{A}}}_{jt\Delta k}$, with timestamp updates Δk of less than 25ms for order book level j for the day t. The total number of messages for order book level j for the day t is $\tilde{\mathcal{M}}_{jt}$. Therefore, the estimated fraction of algorithmic trading messages $\tilde{\mathcal{A}}$ is defined as

$$\tilde{\mathcal{A}}_{jt\Delta k} = \tilde{\mathcal{M}}^{\tilde{\mathcal{A}}}_{jt\Delta k} / \tilde{\mathcal{M}}_{jt} \tag{3.61}$$

The total messages per minute, $\tilde{\mathcal{M}}^{\Delta}_{jt}$, can be calculated as $\tilde{\mathcal{M}}^{\Delta}_{jt} = \tilde{\mathcal{M}}_{jt}/T^*_t$, where

trading time, T_t^*, is the length of the first order to the last order on an individual day t, which is measured in minutes.

3.5 Data and Summary Statistics

Besides the insides quotes dataset I have constructed in the Chapter 2, this chapter utilizes the limited order book data of Eurodollar futures. I consider the order book data from level 1 to level 5 for the 40 quarterly Eurodollar Futures contracts between January 1, 2008 and July 31, 2014. The futures data are labelled with RIC codes, such as GEH0, GEH1,...GEZ9. I provide real-time order book data snapshots in Table 3.1, which displays the first 2 levels of the GEH0 market depth data over 2 sec around 6 pm GMT on August 5, 2008 with milliseconds stamps. The data structure includes the bid prices, bid volumes, the number of buyers, ask prices, ask volumes and the number of buyers in each level.

Table 3.2 indicates the maturity, the average volume and the observations of market depth asks and bids. GE?4 has the largest bid and ask volume with $631.86 quadrillion and $618.16 quadrillion, respectively. The largest average number of observations is the GE?5, which has 57.13 million observations for the bids side and asks side. The market depth data is utilized to compute the liquidity measures and AT proxy for all the five levels. Then I match the daily wavelet variance ratio with the daily liquidity measures and the daily fraction of algorithmic trading. In total, there are 2,339 days with all 40 contracts for my final daily frequency regression analysis.

For the market depth data (i.e. the complete set of quotes within the order-book) we can see a very clear picture of the activity within the market graphically.

Table 3.1: Order Book Data Structure

#RIC	Date[G]	Time[G]	GMT Offset	Type	L1 Bid Price	L1 Bid Size	L1 Buy No	L1 Ask Price	L1 Ask Size	L1 Sell No	L2 Bid Price	L2 Bid Size	L2 Buy No	L2 Ask Price	L2 Ask Size	L2 Sell No
GEH0	05-Aug-08	18:14:19.682	-5	Market Depth	95.885	72	4	95.89	30	1	95.88	25	1	95.895	119	6
GEH0	05-Aug-08	18:14:19.731	-5	Market Depth	95.885	72	4	95.89	30	1	95.88	25	1	95.895	119	6
GEH0	05-Aug-08	18:14:19.731	-5	Market Depth	95.885	119	6	95.895	119	6	95.88	25	1	95.9	25	1
GEH0	05-Aug-08	18:14:19.868	-5	Market Depth	95.885	119	6	95.895	119	6	95.88	25	1	95.9	25	1
GEH0	05-Aug-08	18:14:19.868	-5	Market Depth	95.89	25	1	95.895	99	5	95.88	25	1	95.9	25	1
GEH0	05-Aug-08	18:14:19.960	-5	Market Depth	95.89	40	2	95.895	99	5	95.885	79	4	95.9	39	2
GEH0	05-Aug-08	18:14:19.990	-5	Market Depth	95.89	40	2	95.895	74	4	95.885	79	4	95.9	64	3
GEH0	05-Aug-08	18:14:20.050	-5	Market Depth	95.89	119	6	95.895	15	1	95.88	25	1	95.9	123	6
GEH0	05-Aug-08	18:14:20.050	-5	Market Depth	95.89	119	6	95.9	138	7	95.88	25	1	95.91	5	1
GEH0	05-Aug-08	18:14:20.200	-5	Market Depth	95.89	119	6	95.9	138	7	95.885	25	1	95.915	10	2

Notes: This table shows the first 2 levels of the order book data (market depth) structure of GEH0 quotes on August 5, 2008. I gather Eurodollar futures contract market depth quotes (the first 5 levels, here only display the level one and level two) and Trades, including bid price, bid volume and number of buyers, ask price, ask volume and number of sellers, and trades price and volume in milliseconds timestamp. The data are sourced from the Thomson Reuters Tick History database (August 2014).

Table 3.2: Size Data Sample for Order Book Data

RIC	Maturity	Asks		Bids	
		Average Volume ($ quadrillion)	Average Number of Obs (million)	Average Volume ($ quadrillion)	Average Number of Obs (million)
GE?0	2010	58.75	20.46	58.58	20.46
GE?1	2011	139.01	32.07	137.53	32.07
GE?2	2012	156.87	42.73	155.26	42.73
GE?3	2013	326.52	39.07	316.77	39.07
GE?4	2014	631.86	47.52	618.16	47.52
GE?5	2015	408.38	57.13	409.42	57.13
GE?6	2016	149.00	45.21	149.93	45.21
GE?7	2017	32.85	22.64	32.88	22.64
GE?8	2008	5.51	9.97	5.50	9.97
GE?9	2009	7.34	12.56	7.18	12.56

Notes: This table depicts data sample size for limited order book data. The roll-over years of the 40 Eurodollar quarterly trade futures and the average number of observations plus the average volume of bids and asks. The limited order book data are from the Thomson Reuters Tick History database from July 1, 2008 to January 1, 2014, labelled by the Reuters Instrument Code (RIC; GEH0, GEH1, ..., GEZ9). The exchange ticker is GE, which is the RIC for the globex only trades and quotes. I calculate the average value for GE contracts having their maturity date in the same year, denoted GE?0, GE?1, ..., GE?9. Hence, GE?0 includes four 10-year future contracts from 2000 to 2010 – namely, GEH0, GEM0, GEU0 and GEZ0. Note, that the minimum tick size on the exchange is 1/4 of a basis point for the nearest expiring contract and 1/2 otherwise. This yields 2,339 days times 40 contracts for my final daily frequency regression analysis. Note that I report the total quoted depth of the market $\sum_t \sum_i P_{i,k} V_{i,k}$ for each delivery year across the four quarterly deliveries, where t is the daily index, k is the intraday time index and i is contract type index.

Figure 3.2 plots the price, volume and number of active accounts quoting for a single contract again I use the March 2010 maturing contract, EDH0, as my exemplar. As the limit order book is only from electronic trades the ticker symbol is now 'GEH0' which stands for Globex. For the two randomly chosen days I see that on July 11, 2008, when the contract still has about 20 months to run there were 268,258 updates to the order flow, that is the changes in quoted price (top plot), volume (middle) or active account connections (lower). For the volume and number of traders I have reflected the bids (-ve) about the abscissa axis

to provide a comparison for the volume of bids versus the volume of asks. The different coloured plots represent the levels of the order book (recall this is ordered by price), where blue is level one, green is level two, red is level three, cyan is level four and magenta is level five.

As EDH0 approaches maturity the volume of messages in the order book increases substantially, as does the volume. In Figure 3.3, by August 6, 2009 the number of message updates over the day measured (UTC time) increases by a factor of ten to 2,018,494. Interestingly, whilst more than five levels are available to quote at (indeed for contracts such as the S&P 500 E-mini, I see up to ten layers of quotes in practice) the tapes indicate that no more than five levels are quotes. It is also striking that the price differentials of the quotes are now no more than the minimum tick-size, so at any given point the range from the level five bid to the ask covers 5 cents. One point to note is that whilst the number of active accounts on August 6, 2009 is higher at the peak, it is still under 300.

CME group has conducted its own survey of its traded futures contracts, the advantage that CME has in this respect is that each account connected to the Globex platform needs to declare the type of connection, that is whether the access is for an individual trading account with named traders or to an automated trading platform. Their 2010 report indicates that about 64% of message updates within the limit order book comes from automated trading accounts for the Eurodollar (the fifth most active by messaging after EuroFX, S&P 500 E-mini, Crude Oil and ten-year treasury note futures and the fourth most active by volume). I find that thresholding the update speed at 25ms returns a modal value at around 65%. Time-series variation for my AT proxy by thresholding at 200 to 50ms provides very similar patterns and I explore this choice in more detail later in this chapter.

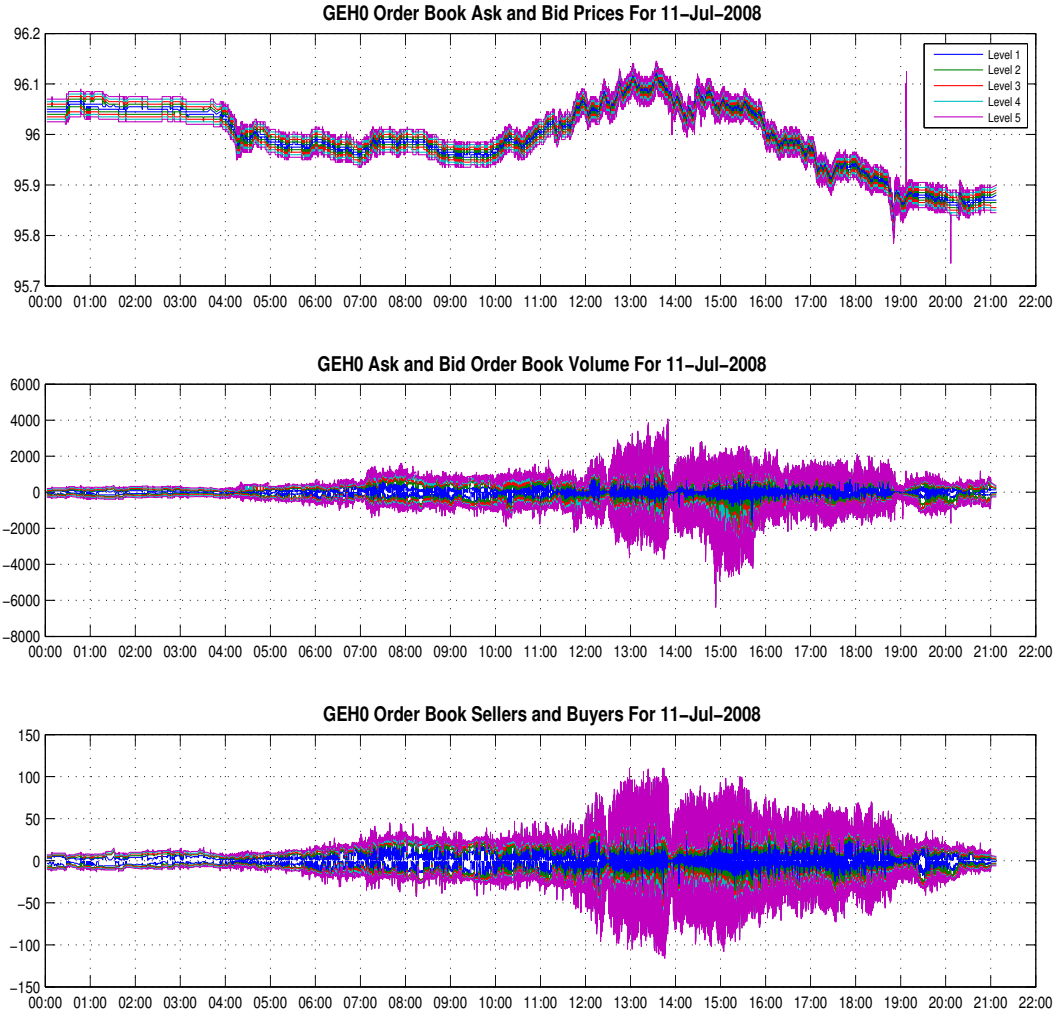


Figure 3.2: Order Book Market Depth in March 2010 Maturing Contract (EDH0 on July 11, 2008)

Notes: This figure displays the five levels limit order book of EDH0 on July 11, 2008: the bid and ask price, volume and the number of buyers and sellers at the five levels. The total volume on this day is 268,258 observations for each level. Notice that whilst the contracts are continuously traded from Sunday to -Friday 5:00 p.m. - 4:00 p.m central standard time (10 pm to 9 pm universal standard time, my plots are all in UCT), there is a 45 minute daily break in trading at 4.15 pm CST (9.15 pm UCT) on weekdays. However, this is not the time at which the daily settlements are computed, this occurs between 1.59 pm and 2.00 pm CST and either reports the mid price of the volume weighted best-bid best-ask or the volume weighted average price of any trades during this survey. For data purposes I treat a day as being 00:00:00.001 CST to 11:59.59.999 CST, although I have run my study by UCT and found no significant difference in the results as the majority of trading is within the day denoted by 00:00:00.001 CST to 11:59.59.999 UCT.

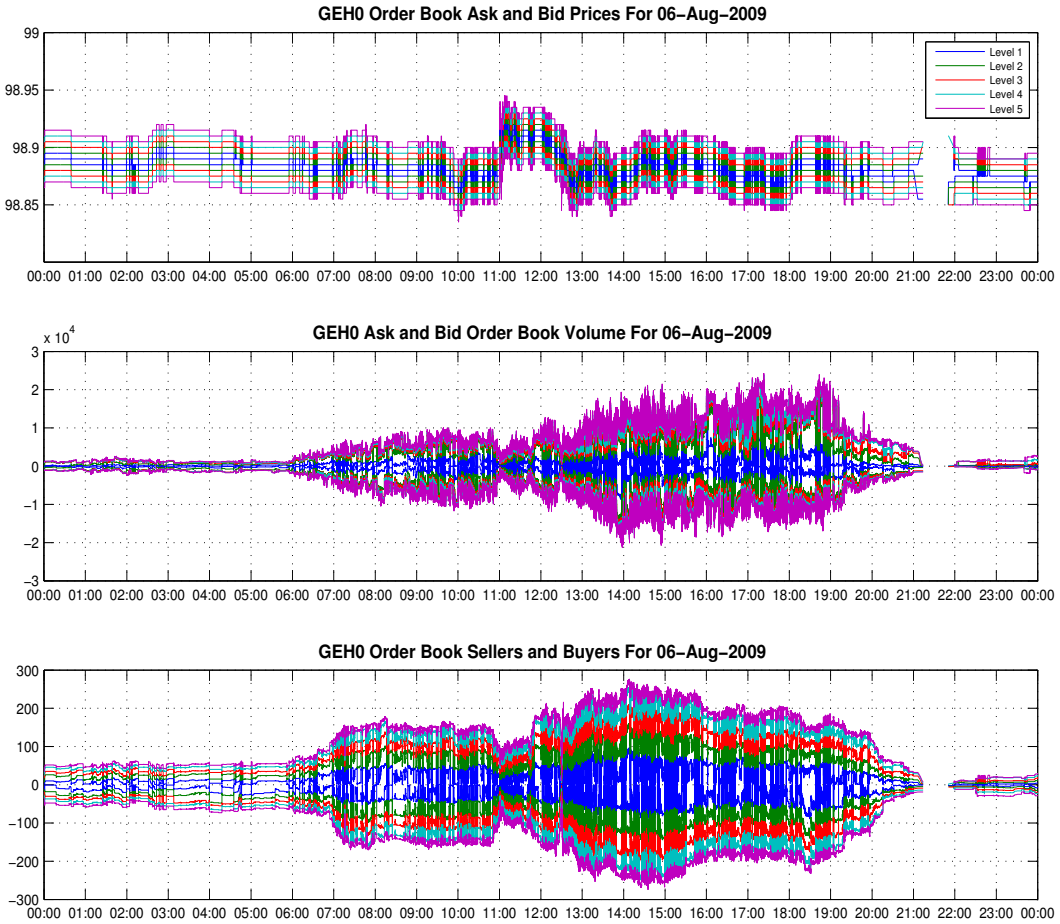


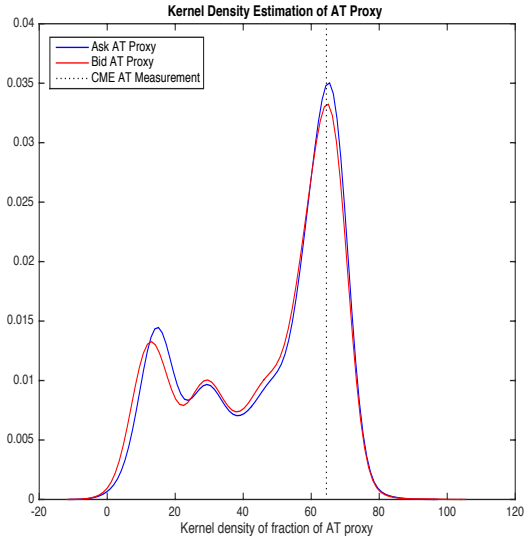
Figure 3.3: Order Book Market Depth in March 2010 Maturing Contract (EDH0 on August 06, 2009)

Notes: This Figure displays the five levels limit order book of EDH0 on August 06, 2009: the bid and ask price, volume and the number of buyers and sellers at the five levels. This is the largest quoting day with 2,018,494 observations for each level. Notice that whilst the contracts are continuously traded from Sunday to -Friday 5:00 p.m. -4:00 p.m central standard time (10 pm to 9 pm universal standard time, my plots are all in UCT), there is a 45 minute daily break in trading at 4.15 pm CST (9.15 pm UCT) on weekdays. However, this is not the time at which the daily settlements are computed, this occurs between 1.59 pm and 2.00 pm CST and either reports the mid price of the volume weighted best-bid best-ask or the volume weighted average price of any trades during this survey. For data purposes I treat a day as being 00:00:00.001 CST to 11:59.59.999 CST, although I have run my study by UCT and found no significant difference in the results as the majority of trading is within the day denoted by 00:00:00.001 CST to 11:59.59.999 UCT.

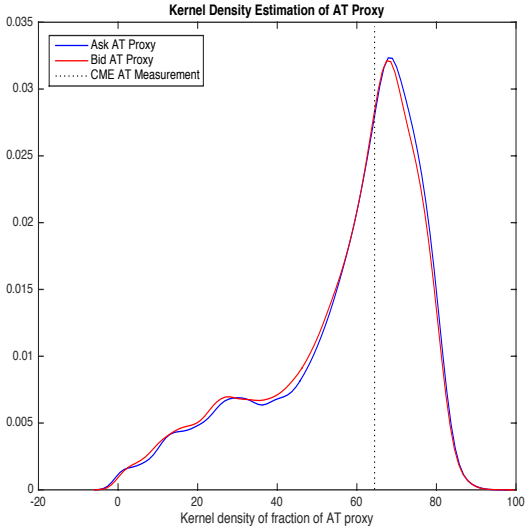
Using the normal kernel function to estimate densities, Figure 3.4 (a) and (b) illustrate the distribution of bid and ask algorithmic trading proxies (threshold at 25 ms) of both the contract-day panel dataset and the volume weighted average dataset. The volume weighted average dataset includes 2,339 observations. When both bid and ask AT proxies are below 5%, the density keeps at a very low level (see Figure 3.4 (a)). Once the fraction of the bid and ask AT reach 14% and 15%, the density of AT proxies suddenly increases to around 0.013 - 0.015, respectively. There are fluctuations for both ask and bid AT proxies between 16% and 40%, where bid AT proxy shows more volatile than ask one. When the fraction of AT traders is larger than 40% of total traders, bid and ask proxies seem to converge and rise steeply, and both arrive the high peak with 0.033 density (bid AT proxy) and 0.035 density (ask AT proxy) around 65%. CME Group [2010] reports that the proportion of algorithmic activities on Eurodollar futures markets is 64.46%, which is consistent with my findings (see the black dotted line in Figure 3.4 (a) and (b)). Finally, AT proxies return to zero at 100%.

Figure 3.4 (b) presents the kernel density of contract-day panel dataset AT proxies with 35,491 daily observations. Consistent with volume weighted average dataset, the faction of AT traders below 5% is very low. Both bid and ask AT proxies follow the similar path and rise smoothly. There are some shifts but the volatile degree is much less than the volume weighted average dataset between 10% and 40%. After the proportion of algorithmic traders larger than 45%, the density of AT proxy suddenly rises and reaches the highest point around 70% and drop back to zero at 100%.

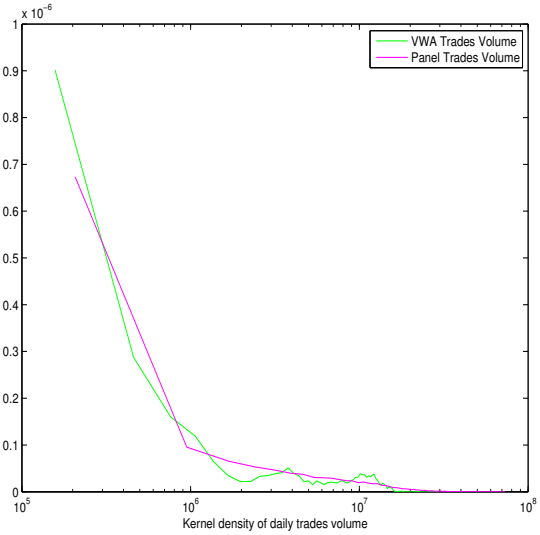
In short, bid and ask AT proxies of two datasets show the similar pattern, with low density level below 5%, the highest density between 65% - 70% and are highly volatile between 15% and 40%.



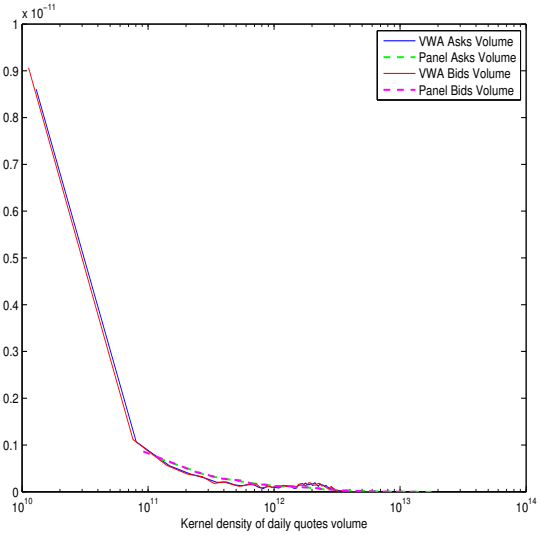
(a) Volume weighted average



(b) Panel



(c) Trades volume distribution



(d) Quotes volume distribution

Figure 3.4: Kernel Density Estimation of AT Proxies and Total Volumes

Notes: Graph (a) shows the kernel density of the bids and asks daily fraction of algorithmic traders with 2,339 observations in the volume weighted average dataset; and graph (b) is the density of daily AT proxies with 35,491 observations in the panel dataset. The black dotted line represents the proportion of AT proxy from the CME Group [2010], which they report that the proportion of algorithmic activities is 64.46% on Eurodollar futures markets. Graph (c) and graph (d) present the kernel density of the trades and quotes daily total volumes in both volume weighted average dataset and panel dataset. The daily total volume is computed as $\sum P_k V_k$, where P_k is trades, asks or bids prices, V_k is trades, asks or bids volume and k is tick-times.

The distribution of the contract volume provides some insight into the activity of the market at a daily frequency; for each day, I compute $\sum V_k$ where P_k are the traded or quoted prices. Figure 3.4 (c) and (d) present the Gaussian kernel density plots (with logarithmic abscissa values) across the contract-day panel. We can see that there are a small number of days in both the panel for trades and quotes that have intense spikes in volume for both the panel and the market aggregate. Unlike the asymmetries in the quoting frequency, however, the distribution of bids and asks is roughly consistent (recall that the volume quoted need not be identical across the bids and asks). The most active days have around $6.65 trillion of contracts traded and roughly three times this quoted.

3.6 Liquidity and Depth Measures with Algorithmic Trading Proxies

At this stage of the chapter, I move from the inside quotes to the complete order book to exploit the comprehensive nature of my dataset. Therefore, from this point of the chapter, the term 'level' refers to one of the five levels of the order book. Latter in the chapter, I study the marginal effects of AT in the quality of the market measured by execution risks. Before that I believe it is important to provide a global view of liquidity, which is a determinant of execution risk, and algorithmic trading in the market. In this context, I first analyze where in the book liquidity is created and how this liquidity evolves over the maturity of the contracts. I summarize my findings in Figure 3.5, where I plot the weekly average of the liquidity measures calculated using the information on every contract in my sample. In general, the patterns of the different liquidity measures are consistent with each other. Both the bid-ask spreads and quoted half spreads are

quite volatile and slowly get narrower and more consistent over time among the five levels. Near to two years before maturity, quoted spreads become narrower indicating greater market liquidity. Then, one week before settlement liquidity suddenly drops again. In terms of quoted depth, there is a clear increasing trend from six years to maturity that is broken only few days before maturity. Following my results it is possible to conclude that the largest liquidity in the market is provided in level two which exhibits the highest depth during these six years and also decreases one week before maturity.

The maturity effect behavior can be clearly seen in the plots of effective half spread, realized half spreads (5 mins and 30 mins) and adverse selection indicators (5 mins and 30 mins). For all levels, Figure 3.5 shows a monotonic decreasing trend during the last five years of the life-cycle of contracts. The narrowest effective half spreads, realized half spreads and adverse selection indicators are observed in level one of the order book. It is worth noting that level one and level two tend to converge, which again highlights the importance of level two.

I now explore the characteristics of AT in the market. In particular, I want to provide a general idea on how much AT activity is observed in the Eurodollar Future market, where in the book the ATs sit (i.e. are they more active on the ask than in the bid or vice versa, do they trade more on a level than other, etc). Table 3.3 presents some descriptive statistics on the order book updates I used to compute the AT proxy (for reference these can be compared with the timestamps of variance ratios from the inside spread data in Table 2.7). The Table shows that although the 200ms limit is lower than the average human reaction time reported by literature on physiology, it is not really a collection point for AT because on the average message updates happen much quicker. In fact, the majority of quotes under 200ms are also under 100ms.

Table 3.3: Order Book Update Frequency

		Order Book Level 1				Order Book Level 2				Order Book Level 3				Order Book Level 4				Order Book Level 5				All Order Book Levels			
		Mean (ms)	Median (ms)	Mode (ms)	Percent (%)	Mean (ms)	Median (ms)	Mode (ms)	Percent (%)	Mean (ms)	Median (ms)	Mode (ms)	Percent (%)	Mean (ms)	Median (ms)	Mode (ms)	Percent (%)	Mean (ms)	Median (ms)	Mode (ms)	Percent (%)	Mean (ms)	Median (ms)	Mode (ms)	Percent (%)
Asks	AT 200ms	33.17	10.94	4.00	69.01	37.22	13.67	5.00	66.45	33.57	10.53	5.00	81.21	23.63	5.19	2.69	86.47	20.89	4.66	2.56	86.37	35.49	12.71	5.00	77.90
	AT 150ms	26.85	10.04	4.00	66.11	30.10	12.53	5.00	63.31	27.33	9.73	5.00	78.49	19.33	4.99	2.69	84.71	17.33	4.54	2.56	84.78	28.69	11.75	5.00	75.48
	AT 100ms	20.36	8.94	4.00	62.07	22.70	10.99	5.00	58.87	20.75	8.59	5.00	74.52	14.89	4.69	2.75	82.15	13.55	4.35	2.56	82.47	21.69	10.33	5.00	72.02
	AT 75ms	16.68	8.16	4.00	58.96	18.49	9.94	5.00	55.47	16.97	7.86	5.00	71.37	12.41	4.44	2.75	80.20	11.45	4.21	2.56	80.69	17.70	9.28	5.00	69.34
	AT 50ms	12.69	7.24	4.40	54.54	13.94	8.48	5.00	50.61	12.81	6.99	5.00	66.64	9.57	4.08	2.75	77.23	9.03	4.00	2.56	77.96	13.38	8.03	5.00	65.40
	AT 25ms	8.15	5.77	4.86	46.61	8.73	6.48	5.00	41.98	8.04	5.47	5.00	57.90	6.31	3.55	2.69	71.46	6.17	3.57	3.02	72.56	8.45	6.21	5.00	58.10
Bids	AT 200ms	33.90	11.78	5.00	68.13	37.30	13.86	5.00	66.50	33.76	10.90	5.00	81.14	24.80	5.38	2.87	84.21	21.86	4.90	2.81	84.85	35.66	13.10	5.00	76.97
	AT 150ms	27.42	10.79	5.00	65.20	30.20	12.76	5.00	63.32	27.48	10.05	5.00	78.42	20.25	5.14	2.87	82.44	18.10	4.73	2.81	83.17	28.86	12.10	5.00	74.51
	AT 100ms	20.83	9.53	5.00	61.14	22.83	11.24	5.00	58.81	20.87	8.90	5.00	74.38	15.52	4.90	2.87	79.84	14.10	4.52	2.81	80.71	21.86	10.66	5.00	70.98
	AT 75ms	17.12	8.66	5.00	58.03	18.63	10.18	5.00	55.37	17.08	8.09	5.00	71.19	12.85	4.62	2.87	77.87	11.89	4.34	2.81	78.81	17.87	9.60	5.00	68.25
	AT 50ms	13.04	7.64	5.00	53.55	14.06	8.66	5.00	50.44	12.90	7.16	5.00	66.44	9.90	4.27	2.87	74.86	9.29	4.11	2.81	75.95	13.53	8.26	5.00	64.25
	AT 25ms	8.37	6.05	5.00	45.49	8.82	6.59	5.00	41.68	8.10	5.59	5.00	57.57	6.48	3.63	2.87	69.15	6.28	3.63	3.21	70.42	8.55	6.37	5.00	56.86

Notes: This table describes both the daily average update frequency across 40 contracts and its standard deviation for each order book level and each algorithmic trading proxy. Here AT proxies are computed with different thresholds (200ms, 150ms, 100ms, 75ms, 50ms and 25ms) for both ask side and bid side. The unit of order book update speed is in milliseconds (ms). For the bid and ask side of the order book, I take the daily average value of updating speed below thresholds across all the five levels and across all 40 For the bid and ask side contracts. See Figure 3.6 and Figure 3.7 for more details.

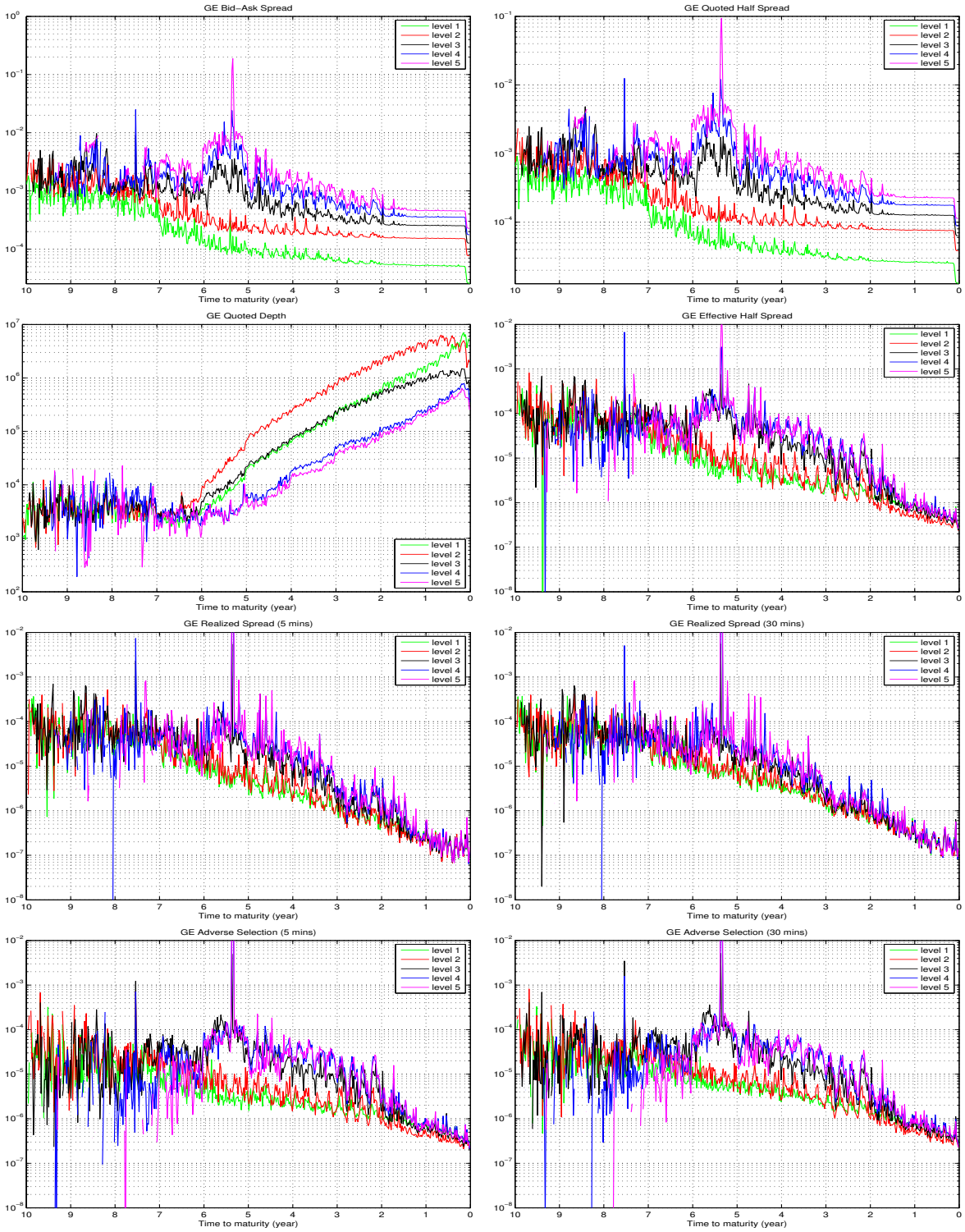


Figure 3.5: Liquidity Spreads

Notes: This figure displays the weekly average of liquidity measures at the five levels for all the contracts including bid-ask spreads, quoted half spreads, quoted depth, effective half spreads, realized spreads (5 mins and 30 mins) and adverse selection (5 mins and 30 mins). Theses spreads are averaged by time to maturity across 40 contracts, hence the x axis is the year to maturity.

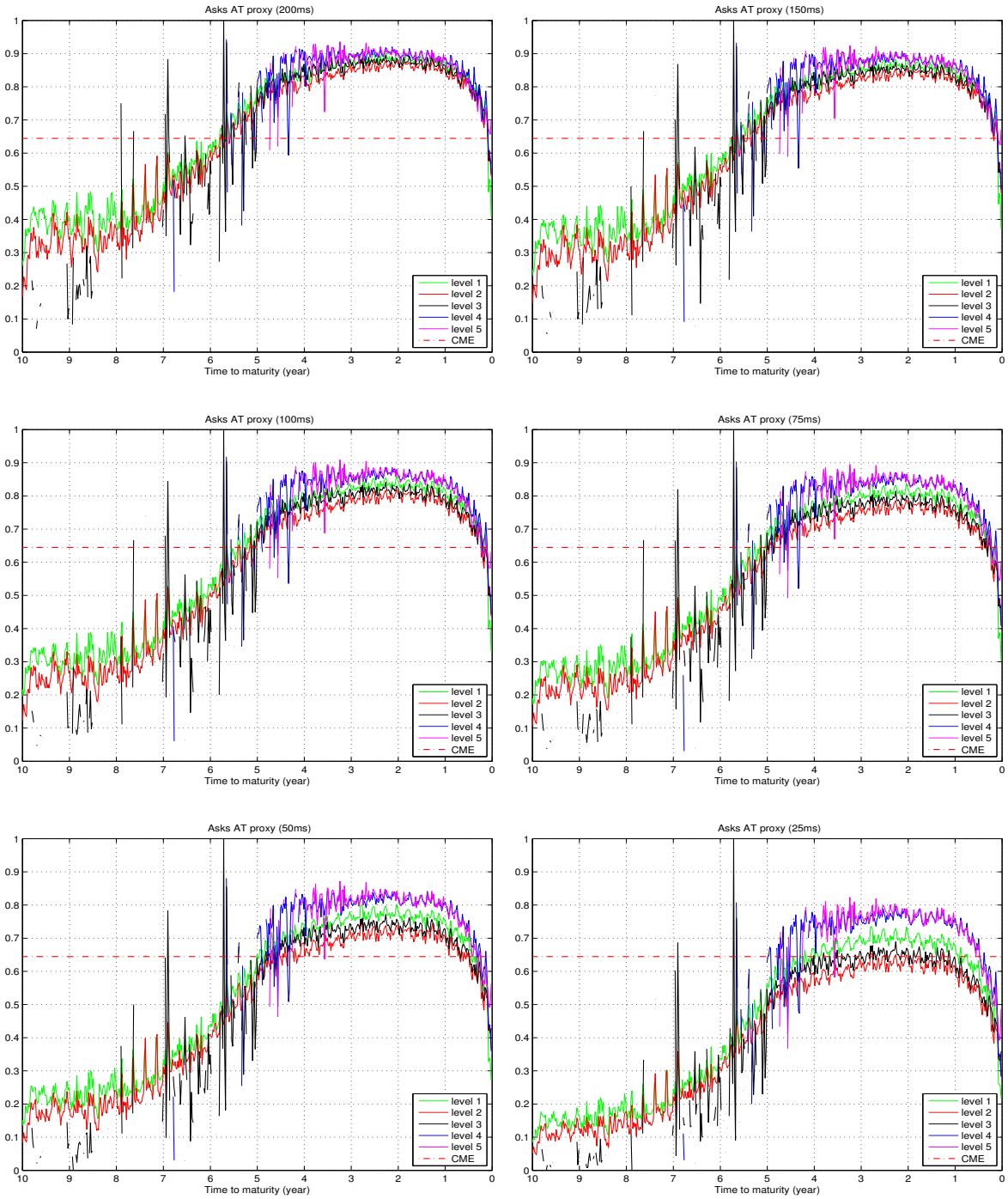


Figure 3.6: Ask Side Algorithmic Trading Fraction

Notes: This figure portrays the weekly average proportion of algorithmic trading in ask sides from level 1 to level 5 across all the contracts. The fraction of algorithmic trading $\tilde{\mathcal{A}}$ is calculated as $\tilde{\mathcal{A}}_{jt\Delta k} = \tilde{\mathcal{M}}^{\tilde{\mathcal{A}}}_{jt\Delta k}/\tilde{\mathcal{M}}_{jt}$, where $\tilde{\mathcal{M}}_{jt}$ is the total number of messages for ask side level j for the day t; and $\tilde{\mathcal{M}}^{\tilde{\mathcal{A}}}_{jt\Delta k}$ is the algorithmic trading messages, defined as a message with time stamp updates Δk of less than the different thresholds (200ms, 150ms, 100ms, 75ms, 50ms and 25ms) for ask level j for the day t. The value of x axis is the years to maturity. The red dashed line represents the proportion of AT proxy from the CME Group [2010], which they report that the proportion of algorithmic activities is 64.46% on Eurodollar futures markets.

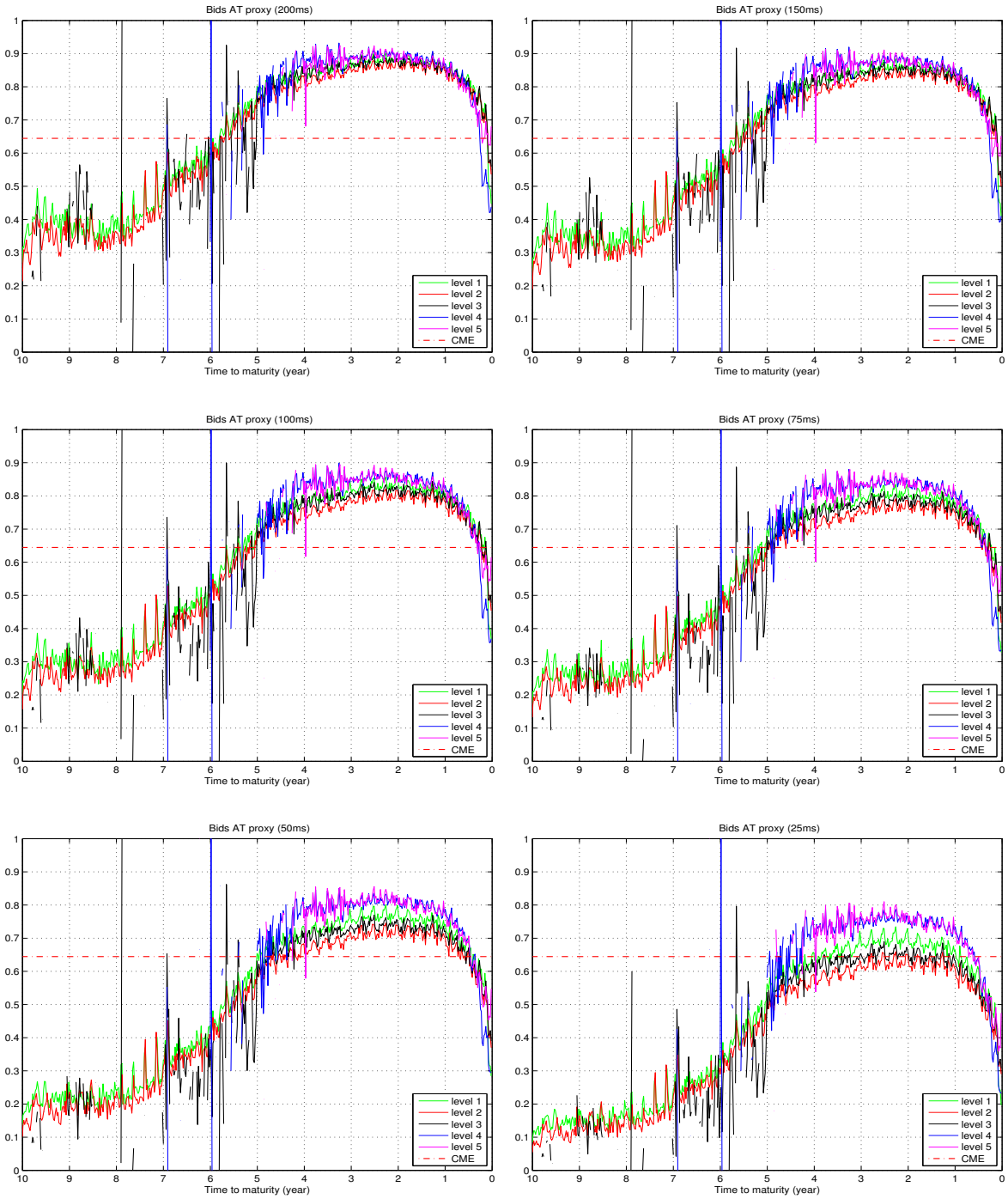


Figure 3.7: Bid Side Algorithmic Trading Fraction

Notes: This figure portrays the weekly average proportion of algorithmic trading in bid sides from level 1 to level 5 across all the contracts. The fraction of algorithmic trading $\tilde{\mathcal{A}}$ is calculated as $\tilde{\mathcal{A}}_{jt\Delta k} = \tilde{\mathcal{M}}^{\tilde{\mathcal{A}}}_{jt\Delta k}/\tilde{\mathcal{M}}_{jt}$, where $\tilde{\mathcal{M}}_{jt}$ is the total number of messages for bid side level j for the day t; and $\tilde{\mathcal{M}}^{\tilde{\mathcal{A}}}_{jt\Delta k}$ is the algorithmic trading messages, defined as a message with time stamp updates Δk of less than the different thresholds (200ms, 150ms, 100ms, 75ms, 50ms and 25ms) for bid side level j for the day t. The value of x axis is the years to maturity. The red dashed line represents the proportion of AT proxy from the CME Group [2010], which they report that the proportion of algorithmic activities is 64.46% on Eurodollar futures markets.

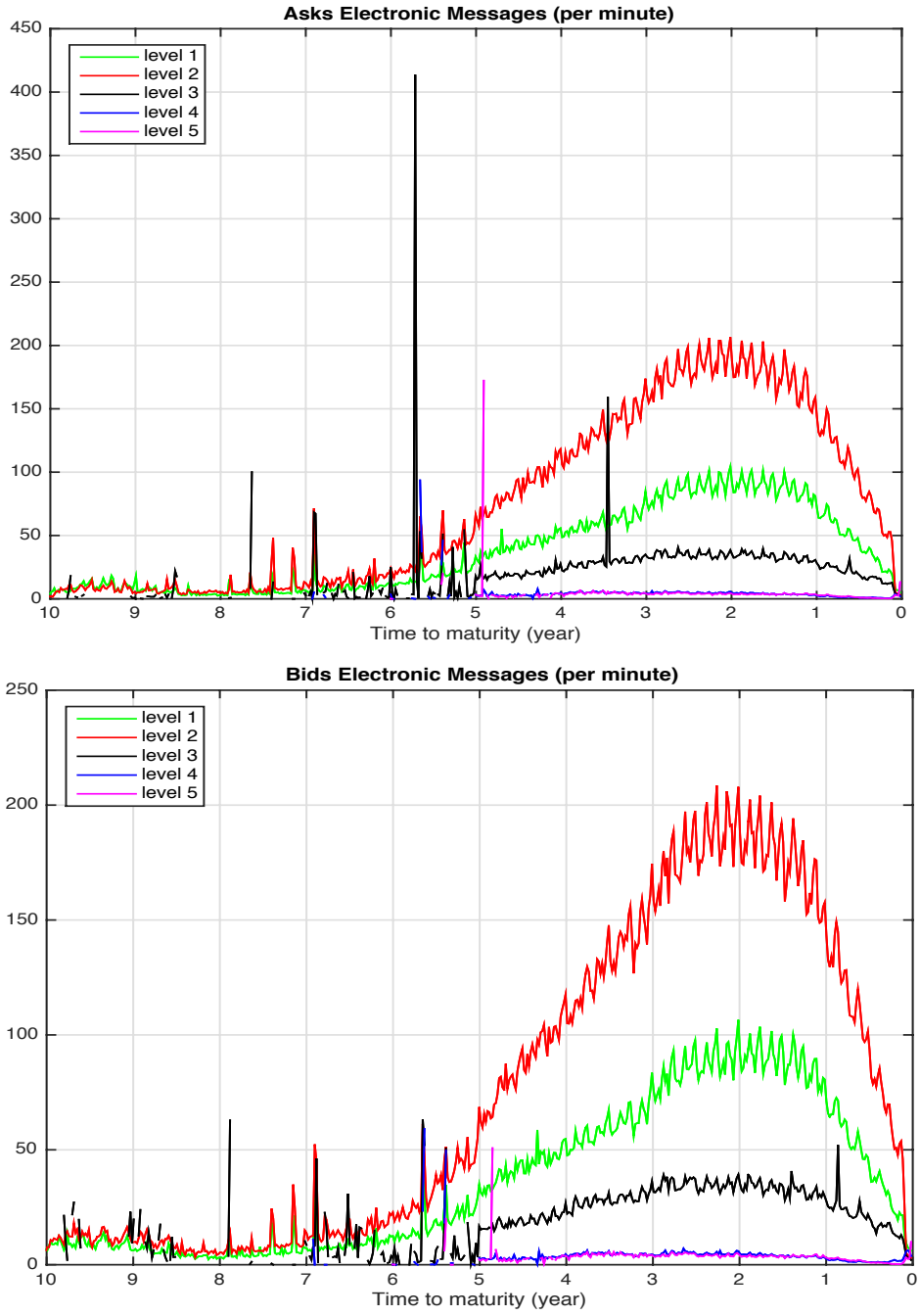


Figure 3.8: Total Messages by Quote Depth Level

Notes: These two graphs capture the weekly average of total messages (per minute) in both ask and bid sides at five different levels. The total messages per minute ($\tilde{\mathcal{M}}^{\Delta}_{jt}$) can be calculated as $\tilde{\mathcal{M}}^{\Delta}_{jt} = \tilde{\mathcal{M}}_{jt}/T^*_t$, where $\tilde{\mathcal{M}}_{jt}$ is the total number of messages for bid or ask side level j for the day t; and T^*_t is the length of the first order to the last order on an individual day t and measured in minutes, j is the order book level. The value of x axis is the years to maturity.

Strikingly, my results also show that censoring at 25ms collects messages that happen on average at 8ms, with most taking place in less than 18ms. This threshold exhibits a median update time between 3ms and 5ms for all levels. The lowest mode update time for all contracts at a 25ms threshold is 2.69ms. Given these results, I turn to a graphical analysis of the activity of AT in each level of the order book. In Figure 3.6 and Figure 3.7, I present the weekly average of algorithmic trading at different thresholds (200ms, 150ms, 100ms, 75ms, 50ms and 25ms) for both the ask and the bid sides.

The fraction of algorithmic trading is quite high for all the five order book levels, even when the AT proxy is set to identify messages at or below 25ms. The chances that with only around 300 active accounts in the market at any given time that the proxy is picking up accidental simultaneous executions seems unlikely as these averages are over an entire day. The variation in the fraction of AT quoting follows a definite term structure and a day of the week effect (Sundays being constantly lower). Considering the 200ms threshold shows that for both the bid and the ask sides the proportion of AT is relatively low (between 20% and 40% for levels 1 and 2 and lower or not existing for the rest of the levels) during the first 2 years of the contracts life-cycle. Then, the proportion of AT increases as maturity approaches, reaching a maximum of around 90% when 2 years remain to maturity. The maximum this time involves all the five levels as AT activity gradually appears in different levels of the order book from 8 to 2 years to maturity. For a 25ms threshold, the maximum proportion of AT is 80% at level 5 and level 4 and 65% at level 2. It is important to note that level 2 is the most rapidly and actively messaged so that as a result it is very consistent with the 64.46% reported by CME Group [2010].

In terms of the total messages, it is observed in Figure 3.8 that shows that level

2 of the order book has the most number of messages per minute. At level 2 for both bid and ask sides, the total messages spike around 200 messages per minute in two years before maturity. What is interesting here is that although level 2 seems to be the predominant one the fraction of the algorithmic trading is still very high at level 4 and level 5 (see Figure 3.8). This means that level 2 is not the major drive for the AT proxy. Therefore, total messages can be treated as an instrument variable to deal with the endogeny effects of AT proxy in the regression, which I will present the estimation results in the next section.

3.7 Chapter Summary

This chapter has provided a theoretical prediction of the marginal effects of the informed traders on the quality of Eurodollar future market, and also conducted the analysis on the high-frequency trading behavior and a broad battery of liquidity provision measurements in this market. I have analyzed the entire market depth of this market from 2008 to 2014 and introduced a new algorithmic trading proxy and battery of tests to assess its impact on the quality of the market. This chapter conducts a population study, in that I have every trade and inside quote from 1996 to 2014 and every message update in the order-book from 2008 to 2014. Importantly, this is the first study to combine both the best-bid best-ask and the order-book data to proxy the influence of high speed quoting on the market. The Globex platform from which most of the orders emanate, does not have direct market order, rather limit orders, market orders with limit protection, market to limit stop limit and stop order with protection, therefore the tapes observe the entire activity of the market over the 2008 to 2014 period, I believe that such a study on the impact of AT on market quality has never before been attempted.

Theoretically, this chapter provides a theoretical prediction that there are non-linear effects of the informed traders on the market quality over the term structure of the Eurodollar contracts. Empirically, I provide a comprehensive view of the liquidity and the algorithmic trading activities in the market using every messaging update at each level within the limit order book from 2008 to 2014.

For the market liquidity, I find the basic bid-ask spread and several others liquidity measures are consistent over the term structure of Eurodollar futures. Liquidity in the initial three to four years is relative stable and low. As the contracts move to the maturity, the liquidity spreads get narrower and more volatiles across all the levels of the order book, especially for those two-year tenor contracts. However, the spreads substantially decrease for the one-week tenor contracts. Besides, my results suggest that the quotes at level two show the highest depth to provide the largest liquidity in the market.

For the high-speed algorithmic trading behavior, this chapter explore that the messaging updating time for the majority of quotes in the limit order book is under 200 milliseconds. It also worth noting that the most active level of the order book for the overall volume of quotes and in particular high speed quoting is level two. Indicating that studies on this contract that ignore quoting outside the best bid best ask are possibly missing some interesting features of the market. My results paint a fascinating picture of the changing behavior of traders in these very actively traded contracts and the impact of algorithmic traders – defined as updating the order-book with update times of under 25ms or 200 ms– on proxies for execution risk and market quality.

Chapter 4

High Frequency Trading and Execution Risk Dynamics: Using an Adapted Semi-parametric Model

Using a novel semi-parametric partial linear model, this chapter assesses the marginal effects of the fraction of high-speed algorithmic traders for execution risks and the quality on Eurodollar future market. I utilize the entire activity within the limit order book and transaction data of the Eurodollar future market from 2008 to 2014. The empirical finding suggests that non-linear relations and certain levels of saturation exist regarding the impact of HFTs on market quality, and this may explain the contradictory evidence found in prior studies of this type. Lower proportions of high-speed algorithmic trading can damage the market quality and increase the execution risks for other traders; higher proportions has the positive marginal effect on the market quality; the impact disappears once the proportions reach certain saturation points.

4.1 Introduction

This chapter introduces a semi-parametric empirical model designed to capture the effects predicted by the theory model from the preceding chapter. I present a new form of the bootstrap estimator based on the original Robinson type estimator with a new bootstrap interval calculation to generate confidence bounds. The semi-parametric estimator has an instrumental variable specification designed to account for endogeneity in the in the proxy I use for high frequency trading (messaging updates under 25 milliseconds or 200 milliseconds). I find that the pattern predicted by the theory model of HFT saturation is supported by empirical evidence over a variety of market quality indicators.

Motivated by the theoretical prediction, I conduct a novel semi-parametric partially linear approach utilizing the fraction of algorithmic trading as an explanatory factor for a series of market characteristics across all 40 Eurodollar quarterly contracts from July 2008 to January 2014. I will demonstrate that: a) that there are critical levels of saturation of algorithmic trading; lower fractions have a detrimental marginal effect on market quality and higher fractions have a positive marginal effect. I believe that this transmission structure may explain many of the contradictory results found in the extant literature. b) I will show that there is a significant term structure effect to the temporal distribution of high frequency trading and that the differential impact on market quality may also go some way to explaining the mixed results on the futures maturity effect conducted on lower frequency data.

I have implemented an adaptation of the conventional semi-parametric partially linear models to account for the instrumental variable approach. To measure the quality of the market, I employ a battery of indicators that make use of a

variety of variance ratios and spreads. For the variance ratios for daily trading in each maturity, I utilize the variance and covariance ratios results from Chapter 2 following the approach of Hasbrouck [2014]. Specifically, the variance-covariance ratio analysis documents the change in the expected level of quadratic variation at different time scales and could be used to determine the impact of high speed quoting on very short term volatility. Additionally, the analysis of the covariance ratio between the quotes on bid and ask side of the inside-spread, provides insight into the time-scale at which there is a breakdown in an agreement between the buy and sell side of the order book, an important measurement of market liquidity and the informativeness of quotes. In supplement to the variance-covariance ratio analysis, I also utilize the battery of liquidity spreads from Chapter 3 implemented in Hendershott et al. [2011], which include the basic bid-ask spread in addition to several others that measure complementary aspects of liquidity and execution risks and costs. I build upon the prior approach of Hendershott et al. [2011] by applying these spread measures to every level of the limit order book.

When I apply the partially linear semi-parametric analysis with the fraction of high frequency trading as an explanatory variable my results are startlingly consistent across measures. When levels of high frequency trading are low (i.e. when the fraction of low frequency traders is higher) the marginal effect of additional high frequency traders reduces market quality across all measures. Interestingly, I find that the level-wise impact on variance ratios as I increase the proportion of high frequency traders is positive (generally considered a deleterious effect), however, the degree of correlation between the bid and ask prices, in contrast decreases (also a deleterious effect). However, this marginal effect completely reverses as the number of high frequency traders passes into the majority. I term this a critical saturation point and postulate that the competition amongst traders results in market outcomes usually deemed as being preferred (higher market liq-

uidity, lower levels of volatility and better agreement between the bid and ask prices). This pattern is also observed across a variety of market indicators.

The reminder of this chapter is organized as follows: Section 4.2 proposes a novel semi-parametric partially linear regression with instruments to capture the relations between the fraction of algorithmic traders and execution risks in the Eurodollar future markets. Section 4.3 provides the analysis of the algorithmic trading at high frequencies and execution risks on Eurodollar future market, whilst Section 4.4 provides some summary comments. For robustness I have conducted a significant number of ancillary treatments and robustness checks, and these are presented in an extensive online appendix.

4.2 A Semi-parametric Model of Market Quality and Algorithmic Trading

My prediction is that the AT proxies influence on the market quality indicators is mostly likely to be a non-linear relation and as previously stated, my AT proxies may exhibit significant endogeneity. I therefore utilize an instrumental variable (IV), partially linear semiparametric regression, to estimate the potentially non-linear relationship between execution risks measures and the level of algorithmic trading from the Eurodollar Futures order book.

Figures 2.5 to 3.7 plot, respectively, the estimated wavelet variance ratios across all 40 contracts sorted by time to maturity for wavelet levels one to nine, my eight different spreads measured by order-book level, the fraction of messages (averaged by the 40 contract types) in the order-book classified by speed of update from 25ms to 200ms (the preferred threshold) for the asks and the same set of plots

for the bid side of the order-book.

The variance ratios and the spreads form the dependent variables in the regression and the AT proxy forms the explanatory variable, where the dependent variables are collected in the vector $Y_{it} = \{\tilde{\mathscr{S}}_{it}^{Q}, \tilde{\mathscr{S}}_{it}^{Q^{1/2}}, \tilde{\mathscr{S}}_{it}^{D}, \tilde{\mathscr{S}}_{it}^{E}, \tilde{\mathscr{S}}_{it}^{R,5}, \tilde{\mathscr{S}}_{it}^{R,30}, \tilde{\mathscr{S}}_{it}^{AS,5}, \tilde{\mathscr{S}}_{it}^{AS,30}, \tilde{\mathscr{R}}_{A1it}, \tilde{\mathscr{R}}_{A4it}, \tilde{\mathscr{R}}_{A9it}, \tilde{\mathscr{R}}_{B1it}, \tilde{\mathscr{R}}_{B4it}, \tilde{\mathscr{R}}_{B9it}, \tilde{\mathscr{R}}_{AB1it}, \tilde{\mathscr{R}}_{AB4it}, \tilde{\mathscr{R}}_{AB9it}\}$. However, as I do not have access to the specific account identifiers for each message update in the order-book I cannot tell for certain whether a message has been created by an algorithmic trader operating at high speed or is just the product of two low-frequency traders coincidently submitting an order. I therefore introduce an instrumentation step, which utilizes the logarithm of total volume of traffic within a day and the logarithm of time to maturity (in days) of the contract as I would anticipate this will be correlated with the dependent variable, but should be orthogonal to the error term if the AT proxy is to be, in part determining the level of execution risk.

The regressions are computed as a contract-day panel across the 40 contracts by day, from July 2008 to January 2014 (the sample period for the order-book data). For my purposes I adapt the time series version of the fractional polynomial estimator for my panel based analysis. The focus of this chapter is not the specification of a new estimator, but to implement a contract-day panel and time series semi-parametric model with instrumental variables with possibly endogenous regressors, I did adapt part of the state-of-the-art in this area to my needs and I shall briefly overview this estimator (as it is not so commonly used in this context) and my approach for generating confidence intervals and marginal effects. The estimator in general terms is expressed as follows:

$$Y_{i,t} = X_{i,t}\beta + \mathscr{G}(Z_{i,t}) + \varepsilon_{i,t}, \quad \mathbb{E}[\varepsilon_{i,t}|W_{i,t}] = 0 \tag{4.1}$$

where $Y_{i,t}$ is my measure of execution risks/market quality for the i contract type (analogous to the ten year term structure of the futures prices) for day t, $X_{i,t}$ are my $e = 2$ linear regressors (in my case the lagged dependent variable and a constant) and $Z_{i,t}$ is my $f = 1$ algorithmic trading proxy. Flattening the data set to an index $i \in \{1, \ldots, N\}$, where N is the total number of day-contracts and the index i represents the set of N tuples of (i, t) contract-days.[1]

Before estimating the semiparametric regression, the endogeneity problems between independent variables X_i (the lagged execution risk measurements) and Z_i (algorithmic trading proxies) with the error terms ε_i need to be concerned. To obtain unbiased estimates, there are some conditions to be satisfied. One is the no multicollinearity between independent variables, $\mathbb{E}[\varepsilon_i | X_i, Z_i] = 0$. The other is the error term should not be correlated with each independent variable, $cov(\varepsilon_i | X_i, Z_i) = 0$. To be more strict, the second assumption should be that the error term should have zero mean conditional on the independent variables, $\mathbb{E}(\varepsilon_i | X_i, Z_i) = 0$.

To address the endogeneity issue, I employ the logarithm of time to maturity and/or the logarithm of total number of messages per day as instrument variables, where denoted as W_i. As valid instrument variables, W_i need to satisfied two conditions: one is the instrument variables W_i should be partially correlated with the endogenous variable Z_i, $cov(W_i, Z_i) \neq 0$. In my case, both the log total number of messages per day and the log time to maturity are correlated with the AT proxy. The other condition requires that $cov(W_i, \varepsilon_i) = 0$. However, this condition can not be tested due to the unobservable of error term ε_i.

[1] A word on notation. I will use mathscript capitals $\mathscr{A}$ to denote functions, vector operators (followed by $:\to$) and densities (which have no following parentheses), $\mathbb{E}[x|y]$ to denote the expectation of a variable x conditioned (potentially non-parametrically) on y, I denotes a suitably sized identity matrix. Lowercase greek letters denote parameter vectors.

I therefore collect my data set into the vectors $Y = (Y_1, ..., Y_i) \in \mathbb{R}$, $X = (X_1, ..., X_i) \in \mathbb{R}^e$ and $Z = (Z_1, ..., Z_i) \in \mathbb{R}^f$. Therefore β is a vector of unknown linear parameters, $\mathscr{G}(.)$ is an unknown nonparametric function, where ε_i is the error term satisfying the condition that $\mathbb{E}[\varepsilon_i|X_i] = 0$ and $\mathbb{E}[\varepsilon_i|Z_i] = 0$. Let $W = (W_1, ..., W_i) \in \mathbb{R}^g$ be the $g = \{1, 2\}$ instruments (in my case the log time to maturity and/or the log total number of messages per day), where $\mathbb{E}[\varepsilon_i|W_i] = 0$. The following presents a brief summary of my semi-parametric estimator and summarize my implementation.

4.2.1 My Adapted IV Robinson Estimator

I adjust the standard instrumental variable semi-parametric partially linear for my purposes. The estimator runs over both panel and time series data, and I have tested a variety of different Kernels with similar results. In my online appendix, I compare my results to the conventional Robinson estimator implemented in Stata 13. One of the major issues with the standard implementations is speed; this approach is usually designed for smaller panel datasets than the one I have deployed here. Furthermore, I need to be able to compare instrumental variables across a number of equations describing the market quality and execution risk and my approach to dealing with this problem is below.

The conditional expectation is unknown as such I utilize the Robinson [1988] double residual methodology by applying a nonparametric conditional mean estimator $\mathbb{E}[Y_i|Z_i]$, $\mathbb{E}[X_i|Z_i]$ and $\mathbb{E}[\varepsilon_i|Z_i]$ on Equation (4.1). I will follow Newey and Powell [2003] and adhere to the exponential family of kernels and more specifically a Gaussian kernel. My major departure from Robinson [1988] and Florens et al. [2012] will be in the bootstrapping procedure I implement to generate the

confidence bounds of the model. To begin with I will outline the family of models, using an adaptation of notation in Florens et al. [2012] and then introduce my bootstrap procedure. Let the general instrumented semi-parametric regression be denoted in expectations as:

$$\mathbb{E}[Y_i|Z_i] = \mathbb{E}[X_i|Z_i]\beta + \mathscr{G}(Z_i) + \mathbb{E}[\varepsilon_i|Z_i] \tag{4.2}$$

By subtracting (4.2) from (4.1), I can build an ordinary least squares (OLS) estimation of model as follows,

$$Y_i - \mathbb{E}[Y_i|Z_i] = (X_i - \mathbb{E}[X_i|Z_i])\beta + (\varepsilon_i - \mathbb{E}[\varepsilon_i|Z_i]) \tag{4.3}$$

where I define $\tilde{Y}_i = Y_i - \mathbb{E}[Y_i|Z_i]$, $\tilde{X}_i = X_i - \mathbb{E}[X_i|Z_i]$ and $\tilde{\varepsilon}_i = \varepsilon_i - \mathbb{E}[\varepsilon_i|Z_i]$. I can recover $\tilde{Y}_i$, $\tilde{X}_i$ from the nonparametric regressions of Y_i and X_i on $\hat{Z}_i$. Then the estimated coefficients vector β can be estimated by Equation (4.3) in the absence of the nonparametric function $\mathscr{G}(Z_i)$. Let β be the standard OLS estimates:

$$\beta = \left(\sum\nolimits_i \tilde{X}_i \tilde{X}_i'^{-1}\right)\left(\sum\nolimits_i \tilde{X}_i \tilde{Y}_i'\right) \tag{4.4}$$

I then regress Z_i onto the filtered values of Y_i to provide an estimate of the non-parametric function $\mathscr{G}(Z_i)$,

$$Y_i - X_i\beta = \mathscr{G}(Z_i) \tag{4.5}$$

However, prior to this estimation stage I need to deal with the problem of the variables X_i and Z_i being endogenous with respect to the disturbance term, as such $\mathbb{E}[\varepsilon_i|X_i, Z_i] \neq 0$, $\mathbb{E}[Y_i|Z_i]$ and $\mathbb{E}[X_i|Z_i]$ cannot be estimated consistently via the non-parametric approach proposed herein.

My approach is in the spirit of Florens et al. [2012] in that I estimate a time series/panel version of the instrumented semi-parametric model with linear covariates specifically. The target function $\mathscr{G}(\cdot)$ and parameter β are the solution of the functional equation

$$\mathbb{E}[Y_i|W_i] = \mathbb{E}[\mathscr{G}(Z_i)|W_i] + \mathbb{E}[X_i\beta|W_i] \tag{4.6}$$

therefore the optimal non-parametric function should be such that the model disturbances are uncorrelated with the time to maturity and/or the total messages per day. As such, following Florens et al. [2012] who in turn derived their approach from Robinson [1988] therefore Equation (4.6) can be rewritten as the following indefinite integrals:

$$\int \mathrm{d}y\ y\ \frac{\mathscr{F}_{Y|W}(y,\cdot)}{\mathscr{F}_W(\cdot)} = \int \mathrm{d}z\ \mathscr{G}(z)\ \frac{\mathscr{F}_{Z|W}(z,\cdot)}{\mathscr{F}_W(\cdot)} + \int \mathrm{d}x\ x\beta\ \frac{\mathscr{F}_{X|W}(x,\cdot)}{\mathscr{F}_W(\cdot)} \tag{4.7}$$

$$\int \mathrm{d}y\ y\ \mathscr{K}_{Y|W}(y) = \int \mathrm{d}z\ \mathscr{G}(z)\ \mathscr{K}_{Z|W}(z) + \int \mathrm{d}x\ x\beta\ \mathscr{K}_{X|W}(x) \tag{4.8}$$

where $\mathscr{F}_{Y|W}$ denotes the joint density of Y and W, and similarly for $\mathscr{F}_{Z|W}$ and $\mathscr{F}_{X|W}$, and $\mathscr{K}_{Y|W}$, $\mathscr{K}_{Z|W}$ and $\mathscr{K}_{X|W}$ indicate the conditional densities of Y_i, X_i and Z_i given W_i respectively. Computing the instrumental version of the partially linear semi-parametric model is non-trivial as both the first and second steps much account for the non-parametric element recalling the objective of othogonalizing the disturbances with respect to the instruments.

I define $L^2_{\mathscr{V}}(\mathbb{R}^f)$ and $L^2_{\mathscr{U}}(\mathbb{R}^g)$ as the Hilbert spaces of square integrable functions with respect to the two densities $\mathscr{V}$ and $\mathscr{U}$. Florens et al. [2012] indicate that $\mathscr{V}$ and $\mathscr{U}$ are the functions $\mathscr{F}_Z$ and $\mathscr{F}_W$ respectively then the approach is identical to Darolles et al. [2011], or if the functions $\mathscr{V}$ and $\mathscr{U}$ are indicator functions then

the approach tends to that of Hall et al. [2005]. Indeed, these two cases can be seen as points at the end of a continuum of model choices the econometrician faces. If $\mathscr{V}$ and $\mathscr{U}$ are identified in the continuous domain $[0, 1]$ I can also recover the approach of Newey and Powell [2003], which is a fully non-parametric approach across the model space.

By multiplying with functions of W, integral Equation (4.8) can be formalised as equation of operators as follows,

$$\mathscr{R} = \mathscr{O}_Z \mathscr{G} + \mathscr{O}_X \beta, \tag{4.9}$$

with

$$\mathscr{R} = \mathbb{E}[Y|W]\mathscr{F}_W/\mathscr{U} \tag{4.10}$$

$$\mathscr{O}_X : \mathbb{R}^e \to L^2_{\mathscr{U}}(\mathbb{R}^g) : \tilde{\beta} \mapsto \mathbb{E}[X'\tilde{\beta}|W]\mathscr{F}_W/\mathscr{U} \tag{4.11}$$

$$\mathscr{O}_Z : L^2_{\mathscr{V}}(\mathbb{R}^f) \to L^2_{\mathscr{U}}(\mathbb{R}^g) : \tilde{\mathscr{G}} \mapsto \mathbb{E}[\tilde{\mathscr{G}}(Z)|W]\mathscr{F}_W/\mathscr{U} \tag{4.12}$$

where $\mathscr{R} \in L^2_{\mathscr{U}}(\mathbb{R}^g)$, $\mathscr{G} \in L^2_{\mathscr{V}}(\mathbb{R}^f)$, $\beta \in \mathbb{R}^e$, and $\mathscr{R} \in \mathcal{R}(\mathscr{O}_X) + \mathcal{R}(\mathscr{O}_Z)$ where $\mathcal{R}(\mathscr{O})$ is the range or co-domain of the operator $\mathscr{O}$. Let $\mathscr{O}_X^*$ and $\mathscr{O}_Z^*$ for $\mathscr{O}_X$ and $\mathscr{O}_Z$, be the following corresponding adjoint operators

$$\mathscr{O}_X^* : L^2_{\mathscr{U}}(\mathbb{R}^g) \to \mathbb{R}^e : \mathscr{J} \mapsto \mathbb{E}[X\,\mathscr{J}(W)] \tag{4.13}$$

$$\mathscr{O}_Z^* : L^2_{\mathscr{U}}(\mathbb{R}^g) \to L^2_{\mathscr{V}}(\mathbb{R}^f) : \mathscr{J} \mapsto \mathbb{E}[\mathscr{J}(W)|Z]\mathscr{F}_Z/\mathscr{V} \tag{4.14}$$

I can identify the operators $\mathscr{O}_X, \mathscr{O}_X^*, \mathscr{O}_Z$ and $\mathscr{O}_Z^*$ using the i.i.d. vectors (Y_i, X_i, Z_i, W_i) from the partially linear model (4.1) in the first step. Recall from Equation (4.9),

the normal equations can be expressed as below,

$$\mathscr{O}_Z^*\mathscr{R} = \mathscr{O}_Z^*\mathscr{O}_Z\mathscr{G} + \mathscr{O}_Z^*\mathscr{O}_X\beta \tag{4.15a}$$

$$\mathscr{O}_X^*\mathscr{R} = \mathscr{O}_X^*\mathscr{O}_Z\mathscr{G} + \mathscr{O}_X^*\mathscr{O}_X\beta \tag{4.15b}$$

Florens et al. [2012] demonstrate that if $\mathscr{O}_X$ is orthogonal to $\mathscr{O}_Z$ (namely $\mathscr{O}_Z^*\mathscr{O}_X = 0$ with $\mathcal{R}(\mathscr{O}_X) \perp \mathcal{R}(\mathscr{O}_Z)$, where $\mathscr{O}_Z^*\mathscr{O}_X$ is a dot product), then I can avoid having to try an iterate estimating $\mathscr{G}(\cdot)$ whilst simultaneously estimating β. So the normal Equation (4.15a) - (4.15b) is equivalent to

$$\mathscr{O}_Z^*\mathscr{R} = \mathscr{O}_Z^*\mathscr{O}_Z\mathscr{G} \tag{4.16a}$$

$$\mathscr{O}_Z^*\mathscr{R} = \mathscr{O}_Z^*\mathscr{O}_Z\mathscr{G} \tag{4.16b}$$

Using kernel estimators to replace the operators $\mathscr{O}_X^*$, $\mathscr{O}_X$and $\mathscr{O}_X^*\mathscr{R}$, yields

$$\begin{aligned} \sum_{i,j} Y_i X_j \frac{\mathscr{K}_h(W_i - W_j)}{\mathscr{U}(W_i)} &= \sum_{i,j} X_i X_j^t \frac{\mathscr{K}_h(W_i - W_j)}{\mathscr{U}(W_i)}\beta \qquad (4.17) \\ \beta &= \left(\sum_{i,j} X_i X_j^t \frac{\mathscr{K}_h(W_i - W_j)}{\mathscr{U}(W_i)}\right)^{-1} \left(\sum_{i,j} Y_i X_j \frac{\mathscr{K}_h(W_i - W_j)}{\mathscr{U}(W_i)}\right) \end{aligned}$$

where the kernel $\mathscr{K}$ is a Gaussian kernel function with bandwidth parameter h ($h > 0$) and where the scaled kernel $\mathscr{K}_h(w) = h^{-g}\mathscr{K}(w/h)$. After identifying the parameter β, estimation of $\mathscr{G}$ can be obtained by purely nonparametric regression [Darolles et al., 2011; Hall et al., 2005].

I can now presume that $\mathcal{R}(\mathscr{O}_X) \not\perp \mathcal{R}(\mathscr{O}_Z)$ with $\mathscr{O}_Z^*\mathscr{O}_X \neq 0$, then so the normal equation (4.15a) - (4.15b) can be expressed as

$$\mathscr{O}_Z^*(I - \mathscr{P}_X)\mathscr{R} = \mathscr{O}_Z^*(I - \mathscr{P}_X)\mathscr{O}_Z\mathscr{G} \tag{4.18a}$$

$$\mathscr{O}_X^*(I-\mathscr{P}_Z)\mathscr{R}=\mathscr{O}_X^*(I-\mathscr{P}_Z)\mathscr{O}_X\beta \tag{4.18b}$$

where $\mathscr{P}_X$ and $\mathscr{P}_Z$ are the orthogonal projections for the $\mathscr{O}_X$ and $\mathscr{O}_Z$, respectively. Hence, $\mathscr{P}_X=\mathscr{O}_X(\mathscr{O}_X^*\mathscr{O}_X)^{-1}\mathscr{O}_X^*$ and $\mathscr{P}_Z=\mathscr{O}_Z(\mathscr{O}_Z^*\mathscr{O}_Z)^{-1}\mathscr{O}_Z^*$. I define the parameter estimators β based on equation (4.18b),

$$\begin{aligned}\beta&=\frac{\mathscr{O}_X^*(I-\mathscr{P}_Z)\mathscr{R}}{\mathscr{O}_X^*(I-\mathscr{P}_Z)\mathscr{O}_X}\\&=\frac{\mathscr{O}_X^*(I-\mathscr{O}_Z(\alpha I+\mathscr{O}_Z^*\mathscr{O}_Z)^{-1}\mathscr{O}_Z^*)\mathscr{R}}{\mathscr{O}_X^*(I-\mathscr{O}_Z(\alpha I+\mathscr{O}_Z^*\mathscr{O}_Z)^{-1}\mathscr{O}_Z^*)\mathscr{O}_X}\end{aligned} \tag{4.19}$$

where α is the positive regularization parameter, which relates to the value of n. I follow the standard approach and assume that the operator $\mathscr{R}$ belongs to $L^2_{\mathscr{U}}(\mathbb{R}^g)$, functions $\mathbb{E}(\mathscr{P}(Z)|W=\cdot)\mathscr{F}_W(\cdot)/\mathscr{U}(\cdot)$ and $\mathbb{E}(X_i|W=\cdot)\mathscr{F}_W(\cdot)/\mathscr{U}(\cdot)$ belong to $L^2_{\mathscr{U}}(\mathbb{R}^g)$ for all $\mathscr{P}\in L^2_{\mathscr{V}}(\mathbb{R}^f)$, and $i\in\{1,...,e\}$. The unknown nonparametric densities estimators (i.e. $\mathscr{O}_X,\mathscr{O}_X^*,\mathscr{O}_Z$ and $\mathscr{O}_Z^*$) can be identified by kernel estimators from the data vectors (Y_i,X_i,Z_i,W_i) from the partially linear model as follows:

$$\mathscr{O}_X\beta = \frac{1}{n}\sum_{i=1}^n X_i\beta\frac{\mathscr{K}_{h_W}(W_i)}{\mathscr{U}} \tag{4.20}$$

$$\mathscr{O}_X^*\mathscr{W} = \frac{1}{n}\sum_{i=1}^n X_i\int\mathscr{K}_{h_W}(W_i-w)\mathscr{W}(w)\mathrm{d}w \tag{4.21}$$

$$\mathscr{O}_Z^*\mathscr{G} = \frac{1}{n}\sum_{i=1}^n \frac{\mathscr{K}_{h_W}(W_i)}{\mathscr{U}}\int\mathscr{K}_{h_Z}(Z_i-z)g(z)\mathrm{d}z \tag{4.22}$$

$$\mathscr{O}_Z^*\mathscr{W} = \frac{1}{n}\sum_{i=1}^n \frac{\mathscr{K}_{h_Z}(Z_i)}{\mathscr{V}}\int\mathscr{K}_{h_W}(W_i-w)\mathscr{W}(w)\mathrm{d}w \tag{4.23}$$

$$\mathscr{R} = \frac{1}{n}\sum_{i=1}^n Y_i\frac{\mathscr{K}_{h_W}(W_i)}{\mathscr{U}} \tag{4.24}$$

where $\mathscr{W}\in L^2_{\mathscr{U}}(\mathbb{R}^g)$, $\mathscr{G}\in L^2_{\mathscr{V}}(\mathbb{R}^f)$ and the bandwidth parameters h_W,h_Z are dependent on sample size N, and the kernel $\mathscr{K}$. The kernels can be any of the

standard kernels, Epanechnikov's, Gaussian or fractional polynomials. For the results presented in this chapter I implemented a gaussian kernel weighted local polynomial fit.

To provide easy to compare results across different models I report the marginal effects of changes in z on y, hence I set $\mathbf{M} = \mathrm{d}y/\mathrm{d}z$ as a function of the dependent variable. If the bootstrapped confidence interval straddles the abscissa axis across its domain then the independent variable has no marginal impact on the dependent variable.

I compute confidence bounds for the partially linear semi-parametric regressions via an i.i.d. bootstrap with 99 resamples. Monte-Carlo studies for this estimator utilizing the function $\tilde{\mathscr{G}} = \mathcal{B}(Z, \mathscr{V}, 2\mathscr{V}) \times \sin(H\mathscr{V}Z)$, where $\mathcal{B}(z, a, b)$ is the probability density function of the beta distribution with shape parameters a, b, the dependent variable is $Z \in [0, 1]$ and the frequency scaling parameter set to $H = 10$.

The bootstrap consistency theory for these types of models is somewhat sparse; however, most implementations use bootstrap to determine the confidence intervals, for single index models a consistency proof for i.i.d. bootstrap is available from Yu and Ruppert [2002].[1] Whilst the choice that I make is the 'third best' approach for providing evidence for the consistency of the bootstrap, it does provide a useful guide across the sets of choices available for my customized implementation versus those in other software packages. I compute the i.i.d. bootstrap in the following steps.

[1] For instance the semipar implementation in Stata uses this approach, I have included some models estimated via a Gaussian Kernel in Stata in the Online Supplement to illustrate that the pattern of the marginal effects is very similar.

Step 1: Estimate the semiparametric regression using the original sample

$$(Y_i, X_i, Z_i, W_i),$$

and compute the fitted value $\hat{Y}_i$ and the residual r_i for each observation.

$$\hat{Y}_i = X_i\beta + \mathscr{G}(Z_i) + \varepsilon_i \tag{4.25}$$

$$r_i = Y_i - \hat{Y}_i \tag{4.26}$$

Step 2: For each fitted value $\hat{Y}_i$, randomly choose a residual r_{ci} to build the bootstrapped Y^* value,

$$Y^* = [\hat{Y}_1 + r_{c1}, \hat{Y}_2 + r_{c2}, ..., \hat{Y}_i + r_{ci}, ..., \hat{Y}_n + r_{cn}] \tag{4.27}$$

where random number $c \in \{1, 2, ..., N\}$. Hence, the bootstrap data sample includes the bootstrapped Y^* value with fitted dependent variables.

Step 3: Refit the model using the bootstrap data sample, (Y_i^*, X_i, Z_i, W_i), and compute the marginal effects $\mathbf{M}$.

Step 4: Repeat the resampling procedure 99 times. The standard deviation of marginal effects $\sigma^{\mathbf{M}}$ can be computed as:

$$\sigma^{\mathbf{M}} = \sqrt{\frac{\sum(\mathbf{M} - \overline{\mathbf{M}})^2}{N-1}} \tag{4.28}$$

Several prior studies, see Florens et al. [2012] for overview have indicated that the distribution of the marginal effects is normal. Therefore, the 95% lower and upper confidence bounds, C^L and C^H, can be calculated by the the 2.5 and 97.5

percentiles of the distribution of **M**.

For each plot, I overlay on a second set of abscissa and ordinate axes the variation of the AT proxy with respect to time to maturity from 10 to 0 years. Inference via confidence bounds is inherently more difficult than via a specific test, such as a standard Wald test, Lagrange multiplier or likelihood ratio. However, my regression framework is relatively simple in construction, with a one contemporaneous non-parametric and one lagged independent variable. For optimal specification of the non-parametric function I use the Akiake Information Criteria (this is used within each bootstrap too hence the variation in the pattern of the confidence bounds).

4.3 Marginal Effects and Execution Risks

In this section, I discuss the marginal effects of algorithmic trading on the market quality indicators that proxy for execution risk utilizing the partially linear semi-parametric model described in Section 4.2. In this framework, execution risk is measured via the variance/covariance ratios and the various of liquidity measures. AT proxies account for the non-parametric component of the regression and the lagged market quality indicators account for the parametric component. Additionally, both the log of total messages and the log of time to maturity are used as instruments to deal with potential endogeneity between the execution risk residual and the algorithmic trading proxy.

On any given day across the forty traded contracts with quarterly delivery, the vast majority of the market. I run regressions on both a 2,339-day market wide average and on 35,491 contract-day panel. The contract-day panel allows me to

instrument the time to maturity directly, whilst the market average I account for possible reverse causality by instrumenting only with the logarithm of total messages for this day. In addition to both datasets, I can instrument for both the volume to message ratio and the total number of traders in the market active at each level within the order book, and report the results of these regressions in the online appendix.

4.3.1 Marginal Effects of AT Proxy Thresholded at 25ms

Table 4.1 records the results of my semi-parametric regressions. The results illustrate the case that, when the AT proxy is regressed against the standard measures of liquidity the coefficients are generally positively and significant. The coefficients attached to AT range from 0.071 to 0.926 for the ask side and from 0.072 to 0.926 for the bid side and, are all significant at a 99% confidence level. When compared, side-by-side, the coefficients on the AT proxy are slightly higher for the bid side when using the volume weighted dataset. This effect is not observed in the panel dataset where there is almost no difference between the explanatory power of AT over execution risk between the bid and the ask sides.

Notice that the R^2 for the panel regression is noticeably higher than the market average. We can see from the sum of evidence in the plots for both execution risk, liquidity and the algorithmic trading proxy that the term structure effect is very strong. Indeed, for any given day the varying effects across the contracts will, to a high level of likelihood, exist and invariably in the most actively traded contracts. As would be expected from a large panel, such as the one to which I have fitted with the single indexed instrument from the first regression are highly significant and I will deal with the non-parametric component forthwith. It is interesting,

Table 4.1: The Semi-parametric Regressions Results with AT Proxy Thresholded at 25ms

		Liqudity Measures								Variance Ratios								
		$\tilde{\mathscr{S}}^{Q}$	$\tilde{\mathscr{S}}^{Q^{1/2}}$	$\tilde{\mathscr{S}}^{D}$	$\tilde{\mathscr{S}}^{E}$	$\tilde{\mathscr{S}}^{R,5}$	$\tilde{\mathscr{S}}^{R,30}$	$\tilde{\mathscr{S}}^{AS,5}$	$\tilde{\mathscr{S}}^{AS,30}$	$\tilde{\mathscr{R}}_{A1}$	$\tilde{\mathscr{R}}_{B1}$	$\tilde{\mathscr{R}}_{AB1}$	$\tilde{\mathscr{R}}_{A4}$	$\tilde{\mathscr{R}}_{B4}$	$\tilde{\mathscr{R}}_{AB4}$	$\tilde{\mathscr{R}}_{A9}$	$\tilde{\mathscr{R}}_{B9}$	$\tilde{\mathscr{R}}_{AB9}$
Volume weighted average																		
	$L(Y_{i,t})$	0.369	0.369	0.926	0.097	0.071	0.578	0.132	0.625	-0.011	-0.001	-0.007	-0.003	0.010	-0.006	-0.003	-0.003	0.002
		(20.67)**	(20.67)**	(137.23)**	(5.43)**	(3.43)**	(34.16)**	(6.47)**	(38.48)**	-0.57	-0.06	-0.32	-0.14	-0.48	-0.30	-0.16	-0.15	-0.09
Ask	$W_{i,t}$	-0.037	-0.018	-0.031	0.003	0.010	0.011	-0.005	-0.005	0.010	0.105	2.856	0.013	0.067	0.390	0.303	-0.864	0.077
AT Proxy		-1.94	-1.94	-1.94	-0.87	-0.74	-0.97	-0.40	-0.46	(3.56)**	(4.03)**	-0.85	(3.46)**	(2.86)**	-1.24	(2.02)*	-1.80	-1.48
	R^2	0.16	0.16	0.89	0.01	0.01	0.33	0.02	0.39	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	N	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339
	$L(Y_{i,t})$	0.382	0.382	0.926	0.099	0.072	0.580	0.133	0.626	-0.010	-0.001	-0.004	-0.003	0.011	-0.004	-0.006	-0.006	0.002
		(20.43)**	(20.43)**	(138.05)**	(4.88)**	(3.51)**	(34.25)**	(6.52)**	(38.60)**	-0.52	-0.03	-0.20	-0.16	-0.54	-0.19	-0.29	-0.31	-0.12
Bid	$W_{i,t}$	-0.042	-0.021	-0.060	-0.006	-0.007	-0.001	0.001	0.001	0.011	0.030	0.760	0.014	0.079	0.644	0.304	-0.922	0.054
AT Proxy		(2.43)*	(2.43)*	(4.32)**	-1.81	-0.63	-0.10	-0.10	-0.07	(4.47)**	-1.35	-0.26	(4.59)**	(3.88)**	(2.36)*	(2.33)*	(2.21)*	-1.18
	R^2	0.15	0.15	0.89	0.01	0.01	0.33	0.02	0.39	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
	N	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339
Panel																		
	$L(Y_{i,t})$	0.735	0.735	0.886	0.656	0.075	0.591	0.131	0.613	0.005	0.046	0	0.003	0.123	0.003	-0.001	-0.001	0.006
		(205.99)**	(205.99)**	(386.41)**	(163.77)**	(14.22)**	(137.96)**	(24.96)**	(146.25)**	-0.87	(8.77)**	-0.07	-0.54	(23.29)**	-0.58	-0.12	-0.12	-1.1
Ask	$W_{i,t}$	0.087	0.043	-0.228	-0.004	0.024	0.02	-0.034	-0.025	0.007	0.017	-3.853	0.006	0.016	-0.02	-0.146	0.465	0.047
AT Proxy		(13.42)**	(13.42)**	(28.59)**	(6.47)**	(2.86)**	(2.67)**	(3.91)**	(3.20)**	(6.63)**	(4.43)**	-1.87	(3.90)**	(4.76)**	-0.09	(2.47)*	(2.45)*	-1.19
	R^2	0.55	0.55	0.83	0.43	0.01	0.35	0.02	0.38	0	0	0	0	0.02	0	0	0	0
	N	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491
	$L(Y_{i,t})$	0.736	0.736	0.886	0.655	0.076	0.591	0.132	0.613	0.005	0.046	0.001	0.003	0.122	0.003	0.000	0.000	0.006
		(206.47)**	(206.47)**	(388.36)**	(163.68)**	(14.27)**	(138.00)**	(25.02)**	(146.29)**	-0.88	(8.76)**	-0.20	-0.54	(23.20)**	-0.50	-0.01	0.00	-1.07
Bid	$W_{i,t}$	0.080	0.040	-0.244	-0.005	0.018	0.009	-0.028	-0.014	0.007	0.013	-0.270	0.006	0.021	-0.047	-0.015	0.039	0.053
AT Proxy		(12.17)**	(12.17)**	(30.15)**	(6.51)**	(2.06)*	-1.21	(3.23)**	-1.75	(6.13)**	(3.32)**	-0.13	(3.97)**	(5.93)**	-0.20	-0.24	-0.20	-1.31
	R^2	0.55	0.55	0.83	0.43	0.01	0.35	0.02	0.38	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
	N	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491

* p <0.05; ** p<0.01

Notes: This table illustrates coefficients estimate results from the semi-parametric regression: $Y_{i,t} = X_{i,t}\beta + \mathscr{G}(Z_{i,t}) + \varepsilon_{i,t}$, where $Y_{i,t}$ is the measure of execution risk/market quality for the i contract type for day t, $X_{i,t}$ are the $k = 2$ linear regressors, $Z_{i,t}$ is the $p = 1$ algorithmic trading proxy thresholded at 25ms, and $W_{i,t}$ is a set of instrumental variables, where $\mathbb{E}[\varepsilon_{i,t}|W_{i,t}] = 0$. Flattening the data set to an index $i \in \{1, \ldots, N\}$, where N is the total number of day-contracts and the index i represents the set of N tuples of (i, t) contract-days. In this table, the dependent variables measuring the execution risk $Y_{i,t}$ are reported into two groups: liquidity measures and variance/covariance ratios. Liquidity measurements contain the bid-ask spreads ($\tilde{\mathscr{S}}^{Q}$), quoted half spreads ($\tilde{\mathscr{S}}^{Q^{1/2}}$), quoted depth ($\tilde{\mathscr{S}}^{D}$), effective spreads ($\tilde{\mathscr{S}}^{E}$), realized spreads 5 minutes ($\tilde{\mathscr{S}}^{R,5}$), realized spreads 30 minutes ($\tilde{\mathscr{S}}^{R,30}$), adverse selection 5 minutes ($\tilde{\mathscr{S}}^{AS,5}$) and adverse selection 30 minutes ($\tilde{\mathscr{S}}^{AS,30}$). Variance/covariance ratios include ask variance ratios ($\tilde{\mathscr{R}}_{A}$), bid variance ratios ($\tilde{\mathscr{R}}_{B}$) and bid-ask covariance ratios ($\tilde{\mathscr{R}}_{AB}$) at level 1, level 4 and level 9. The lagged one dependent variable is denoted as $L(Y_{i,t})$. Instrumental variables are adopted to deal with the endogeneity problems, named as $W_{i,t}$. The volume weighted average dataset uses algorithmic trading proxy instrumented on the log of total messages, and the panel dataset adopts algorithmic trading proxy instrumented on both the log of total messages and the log of time to maturity.

albeit expected, to note that the R^2 for the higher variance/covariance ratios are effectively zero, as much of the signal has already been extracted. Recall, that as I add increasing wavelet levels the local price level signal disappears and the remainder tends towards noise as the sample is more aggressively partitioned and the average extracted; level one only partitions the sample across a single division point.

I now move on to the non-parametric component of the regression. In my online appendix, I present all of the underlying plots of the estimated non-parametric function. However, for comparison herein, I present only the marginal effects data with the variation of the bid and ask AT proxy overlaid in red. In the online appendix, I also present the regressions using the average of bid and the ask AT proxy.

The first point of note, as mentioned in Section 4.2 regression models of this type can suffer from overfitting outside of the bulk of the data, this is so called 'Trimming' problem. The distribution of the AT proxy is, in the main, bounded between approximately 20% and 90% of transactions, see Figure 3.4 and indeed this is the case when choosing any refresh rate from 25 ms to 200 ms, see previous discussion and Figures 3.6 to 3.7 for more details. I see that in several plots the fitted function does exhibit a substantial level of oscillation, however, the bootstrapped confidence bounds appear to correctly identify this as insignificant (subject to the usual caveats regarding confidence bounds versus parametric identification of significance). However, in the main, the non-parametric function does appear to capture a great deal of the complex interactions that would diminish the ability of a linear regression to capture them, or for a certain subsamples result in a misleading indication of a significant linear correspondence. Indeed, my results may provide some insight into why many studies on algorithmic trad-

ing yield diametrically opposite results as it appears that the marginal effect not only changes sign when the AT proxy is increasing; however, the effect of high speed quoting on the bid and ask side is subtly different, I will now explore these results in more detail.

In Figure 4.1 the black line represents the marginal effects of algorithmic trading on the bid-ask spreads. The marginal effects of algorithmic trading on the bid-ask spreads have been measured by two sets of data, one is the volume weighted average dataset in Subplots (a) and (b), the other is the panel dataset in Subplots (c) and (d). The red dashed line shows the proportion of daily AT proxy thresholded at 25 ms for the order book level 2.

Figure 4.1 Subplots (a) to (d) indicate that a small number of algorithmic traders at high frequency have effects on the bid-ask spreads. With the volume weighted average dataset, Subplot (a) and (b) suggest that AT has negative effects on the bid-ask spreads when the AT proxy below 10%. Then the marginal effects on the bid-ask spreads increase to the peak when ask-side AT proxy reaches around 24%. Whereas it drops to the bottom when the AT proxy is across 27% on the ask side, which suggests the algorithmic traders at high frequency provide liquidity into the market, narrow bid-ask spreads, and decrease the trading costs. However, the marginal effect of the AT proxy on the bid-ask spreads suddenly spikes at the AT proxy increasing to 41%. In this period, algorithmic traders at high speeding level act as informed traders to submit and cancel quotes orders, which increase the transact costs and reduce the market liquidity. Once the AT proxy increases over 50%, the ask-side AT has limited effects on the bid-ask spreads. Subplot (b) indicates the effects of bid-side AT proxy on the bid-ask spreads show the similar patterns compared to the ask sides with the volume weighted average dataset. The marginal effects of the bid-side AT on the bid-ask spreads spike around the

24% and 39% proportion and drop to the bottom around 31% proportion.

Figure 4.1 Subplot (c) and (d) present the marginal effects on the bid-ask spreads using the panel dataset. The results suggest that a small proportion of algorithmic traders at high-speeding level have the influence on the bid-ask spread, especially on the bid side; whereas once the AT proxy increases over 40%-45%, the marginal effects on bid-ask spread are limited. Subplot (d) demonstrates that when the AT fraction is between 11% and 18%, the marginal effects of the bid-side AT proxy on the bid-ask spreads decrease dramatically. In this low AT fraction period, high-speeding algorithmic traders act as liquidity providers to reduce the bid-ask spreads and decrease the transaction costs in the market. From 18% to 26%, the marginal effects on the bid-ask spreads suddenly increase, which could be caused by the algorithmic traders acting as the inform traders receiving information earlier than other traders. These front-run high-frequency traders usually dominate quote-matching strategy to submit and cancel a large amount of orders, which increase the trading costs for other traders and reduce market liquidity.

Figure 4.2 represents the marginal effects of algorithmic trading on adverse section (5 mins) in Subplots (a) and (b) and, on adverse selection (30 mins) in Subplots (c) and (d), both with the volume weighted average dataset. This Figure indicates that the increasing AT activities have effects on both adverse selections within both 5 minutes and 30 minutes. The red dashed line shows the proportion of daily AT proxy thresholded at 25 ms for the order book level 2.

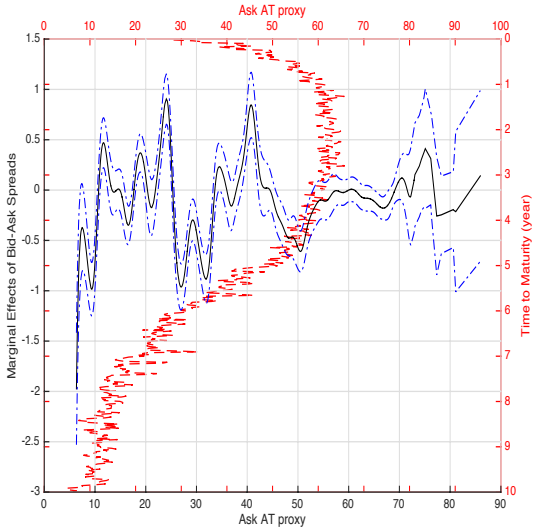
Subplots (a) and (b) suggest that AT has limited effects on the adverse selection (5 mins) when the AT proxy below 40%. Once the AT proxy increases over 40%, the effect of AT on adverse selection (5 mins) rises dramatically to the peak. This indicates that ATs increase trading costs, decrease the market liquidity, and

damage the price efficiency with more informative quotes in the market. At this stage, the increasing algorithmic traders act as informed traders to submit and cancel quotes orders. The degree of asymmetric information in the market reaches a higher level with increased the transaction cost to execute orders. However, the marginal effect suddenly drops as the ask-side AT proxy increases to about 46% - 53%. In this period, ATs provide liquidity to the market and increase price efficiency with less informative quotes.

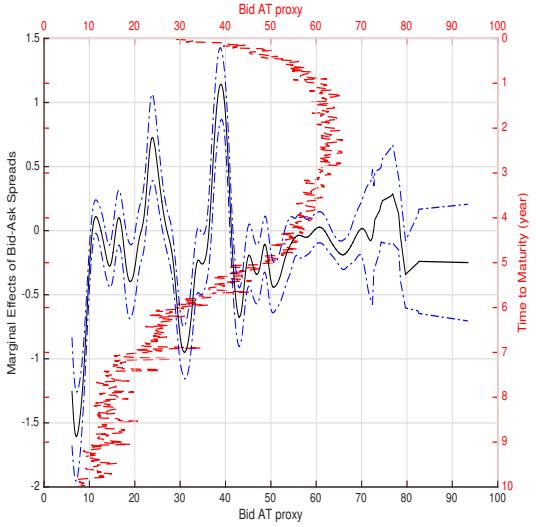
In Subplots (c) and (d), we can see that the bid side AT proxy, the marginal effects are markedly more volatile than those computed from the ask side. There are no distinct effects of ask side AT proxy on the adverse selection (30 mins) when the proportion of algorithmic traders is below around 36%. Within 36% to 44% proportion, the marginal effects are positive. Once the fraction of AT passes 44%, the impact of ask side AT proxy on adverse selection (30 mins) drops until the AT fraction reaching around 55%.

Overall, the proportion of algorithmic trading significantly impacts on the adverse selection (5 mins and 30 mins) and market liquidity. AT shows limited impacts when the AT proxy is relative low. As the increase of AT proportion, the algorithmic traders act as liquidity demanders, which ATs act as informed traders and its impacts on adverse selection increase. Once the AT reaches a certain critical point, the algorithmic traders act as liquidity providers, which leads to the shrinkage of asymmetry information and the increasing market liquidity. When the most traders in the market are ATs, the marginal impacts effectively drop to zero.

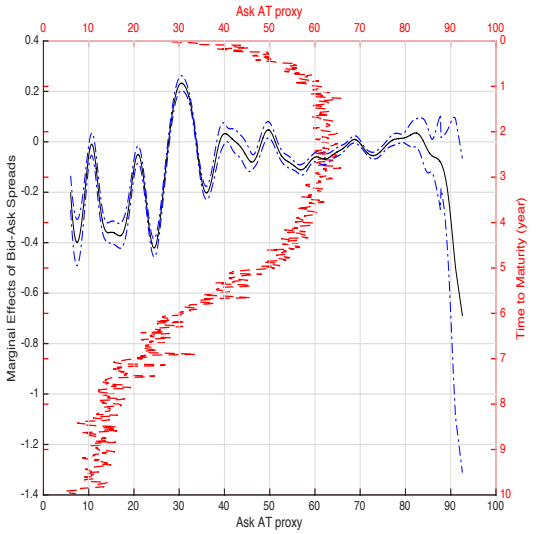
In the vast majority of the cases neither the bid AT proxy nor the ask AT proxy is significant, therefore suggesting that execution risk is not related to the presence of AT in these futures market. A notable exception is the case where execution



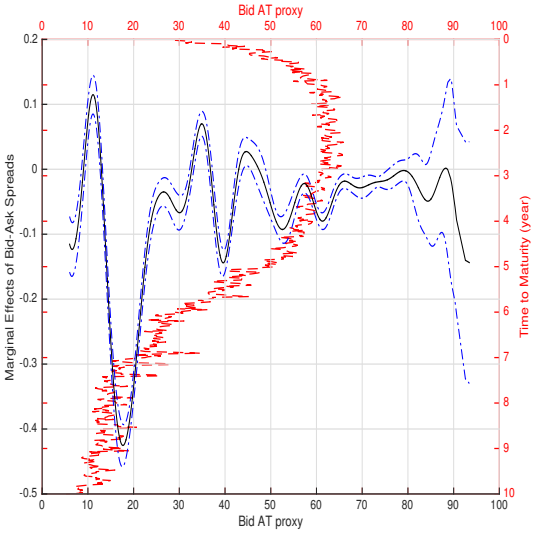
(a) Bid-Ask Spreads - VWA



(b) Bid-Ask Spreads - VWA



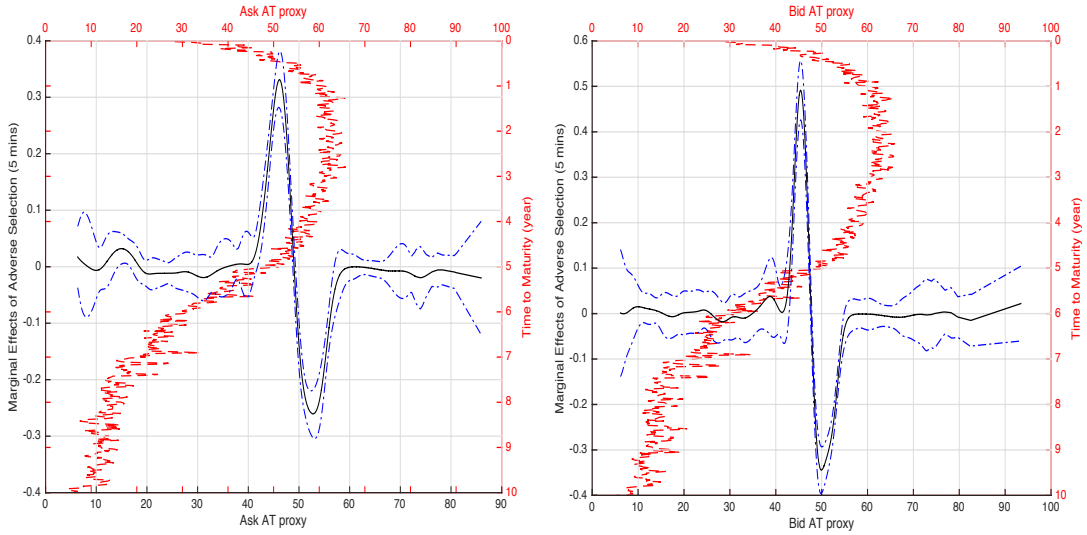
(c) Bid-Ask Spreads - Panel



(d) Bid-Ask Spreads - Panel

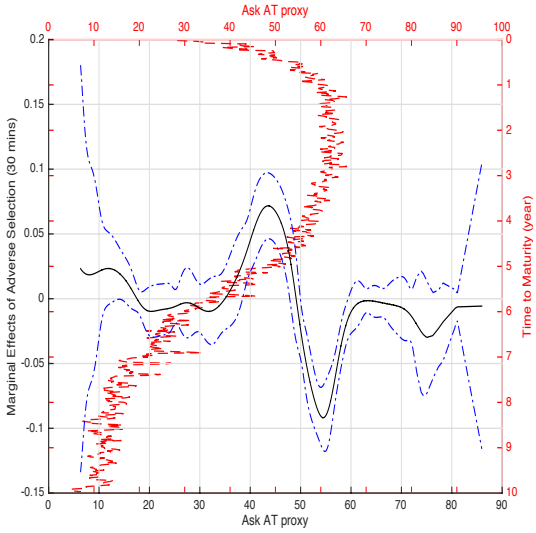
Figure 4.1: Marginal Effects of Bid-Ask Spreads Thresholded at 25 ms

Notes: The red dashed line represents the evolution of daily AT proxy thresholded at 25 ms for order book level 2 in function of the average time to maturity. The black continuous line depicts the marginal effects of AT on the bid-ask spreads in Subplots (a) to (b) with the volume weighted average dataset and Subplots (c) and (d) with the panel dataset. The blue dashed lines are the 95% lower and upper confidence bounds of bid-ask spread marginal effects.

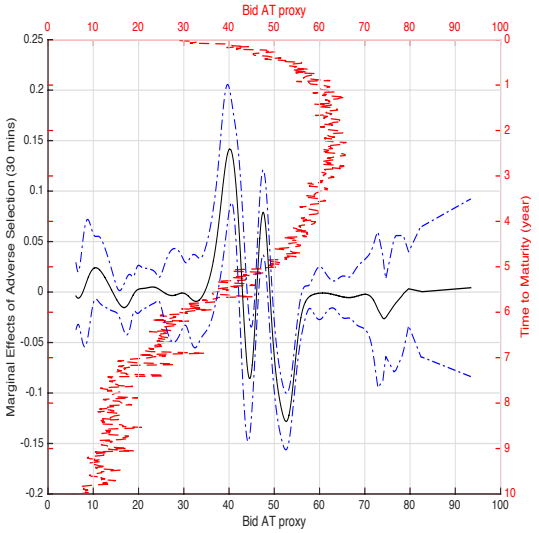


(a) Adverse Selection (5 mins) - VWA

(b) Adverse Selection (5 mins) - VWA



(c) Adverse Selection (30 mins) - VWA



(d) Adverse Selection (30 mins) - VWA

Figure 4.2: Marginal Effects of Adverse Selection (5 mins and 30 mins) Thresholded at 25 ms

Notes: The red dashed line represents the evolution of daily AT proxy thresholded at 25 ms for order book level 2 in function of the average time to maturity. The black continuous line depicts the marginal effects of AT on adverse selection (5 mins) in Subplots (a) to (b) and, on adverse selection (30 mins) in Subplots (c) and (d), both with the volume weighted average dataset. The blue dashed lines are the 95% lower and upper confidence bounds of adverse selection marginal effects.

risk is considered at level 4, therefore at medium frequencies, and defined either as ask variance ratios or as bid variance ratios. In this case, both bid and ask AT have a significant impact on execution risk. Although, in general, the results from variance ratio are not as good as the ones observed using the liquidity measures, the marginal effect analysis shown in Figure 4.3 to Figure 4.5 clearly reveals that the variance-covariance measures tell only a part of the story and cannot be dismissed. See Budish et al. [2013] for materially similar results in a different setting.

Each of these Figures combines the information on how the activity of algorithmic traders evolves with the time to maturity of futures contracts and the information about the marginal effects of algorithmic trading on my variance-covariance measures of execution risk. It is worth to remind that the salient characteristic of these measures is their capacity to isolate the level of decoupling between both sides of the book affecting solely high frequency traders. [1]

Following on from my preliminary analysis which revealed that most of the AT activity happens at the second level in the order book, all the Figures on this section show evidence that level two is critically important in explaining the market characteristics. The information provided by each of these Figures allow me to describe in more detail how the decoupling between both sides of the book depends on the proportion of AT actively placing orders in the market and, how the AT activity evolves over the life-cycle of the futures contract.

In Figure 4.3 the dashed red line represents the evolution of daily AT proxy for order book level 2 in function of the average time to maturity. The black continuous line depicts the marginal effects of AT on variance ratio in Subplots

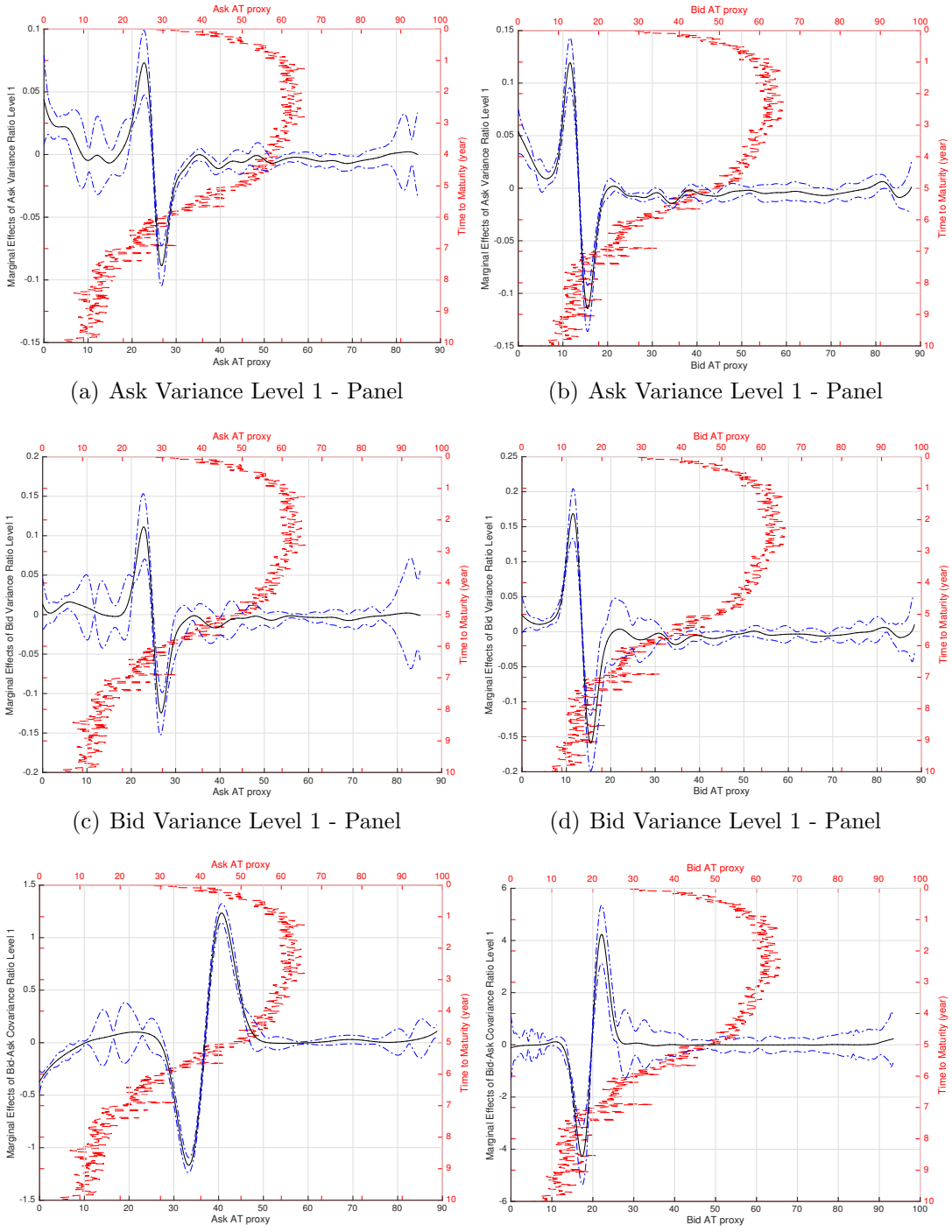
[1] In the thesis, I only report the results of high frequency traders defined as traders placing orders within a threshold equal or lower than 25ms or 200ms. However, Figures would be qualitatively similar for alternative thresholds such as 50ms, 75ms,100ms and 150ms.

(a) to (d) and, on covariance ratio in Subplots (e) to (f), both at timescale 1 with the contract-day panel dataset. The blue dashed lines are the 95% lower and upper confidence bounds of marginal effects. Subplots (a) and (b) suggest that AT marginal effects on ask variance ratio level 1 depend on the proportion of algorithmic trading activities in the market. When this proportion is lower than approximately 20%, the evidence suggests that AT has limited effect on the variance ratio of the spreads. Once this proportion of AT is higher than 20%, there is an increase in the marginal effect on the variance ratio revealing AT increase the high-frequency microstructure noises and volatilities in the market. At the same time, higher variance ratios are symptomatic of less informative quotes and, therefore, higher execution risk as the fundamental price of the contracts is more difficult to observe in the market. However, as the proportion of AT increases further (to around reaches 23 - 25%) the effect on the variance ratio drops dramatically most likely because the high proportion of AT results in their orders cancelling each other, shrinking liquidity and, in turn, the spreads variance.

The marginal effects for the bid variance ratio level 1 in Subplots (c) and (d) present the similar pattern compared with the ask variance ratio. When the ask side AT proxy is below 20%, the impact of algorithmic trading on bid variance ratio is relatively weak. When the proportion of ATs is around 24%, its impacts on bid variance suddenly spike to the top. Once the fraction of ATs is higher than 24%, the marginal effects drop dramatically. Then the marginal effects on the bid variance ratio level one increase, when the ask side AT proxy passes around 27%. Eventually, there are no much effects on the bid variance ratio at level 1, when the ask side AT proxy is higher than 30%. For the impact of the bid side AT proxy on the bid variance ratio at level 1, the pattern is similar to the ask side; however, the point of inflection is lower than for the ask side, at around 10% - 20% of messages being ascribed to be algorithmic in origin.

The evidence suggests that once the AT proxy reaches the critical point 20%, variance ratios will stabilize. I conjecture that this is because once the majority of quoted volume belongs to algorithmic traders using the similar algorithms the arms race has saturated the market as each sends out the similar signals. I define this critical point as AT market saturation, I believe that it is this population measurement that explains the differing results from prior papers in the literature that have only had access to short snippets of the order-flow over say a single month or in some cases only a few days.

Subplots (e) and (f) tell a similar story but in terms of decoupling of the bids and asks. As in the previous case, the effect of AT on the covariance ratio at time scale 1 are virtually non-existent before the proportion of AT reaches around 28% and after it increases further than 48-50% approximately. Within a 28% to 33% proportion the impact of AT is negative as the proportion of AT results in a significant decoupling of bid - ask spreads. Then, up to a proportion of 42% AT substantially increases the covariance ratio, therefore reducing the decoupling; one of key interest here is the shape of the transmission function for the covariance versus the variance, notice that in keeping with the saturation argument the deterioration in market quality (the reduction in covariation evidencing a decoupling) occurs at lower fractions of high speed messages. However, the points of inflection for the bid versus the asks are markedly different and indeed these points differ from the inflection points for the variances. The reduction in covariation of the level 1 bid-ask spread is impacted at a substantially lower fraction of high speed AT messages than for the bid AT proxy. The separation of the volatility and correlation effects at level one is quite fascinating and indicates that the mix strategies employed by HFTs obviously have markedly differentiated impacts.



(a) Ask Variance Level 1 - Panel

(b) Ask Variance Level 1 - Panel

(c) Bid Variance Level 1 - Panel

(d) Bid Variance Level 1 - Panel

(e) Bid Ask Covariance Level 1 - Panel

(f) Bid Ask Covariance Level 1 - Panel

Figure 4.3: Marginal Effects of Variance and Covariance at Timescale Level 1 Thresholded at 25 ms

Notes: The red dashed line represents the evolution of daily AT proxy thresholded at 25 ms for order book level 2 in function of the average time to maturity. The black continuous line depicts the marginal effects of AT on variance ratio in Subplots (a) to (d) and, on covariance ratio in Subplots (e) and (f), both at timescale 1 with the panel dataset. The blue dashed lines are the 95% lower and upper confidence bounds of marginal effects.

Which strategies may cause these effects? Evidence from a variety of legal actions have suggested that HFT strategies broadly fall into two categories, trading across markets, such as spot-future arbitrage or trading within a single market. The obvious causation for excess variance in the spread would be from momentum ignition strategies, whereby the algorithm generates quotes (usually at level 2, so they are no transacted) and uses this volume impact on the order-flow to force the mid-price and inside quotes in a specific direction (e.g. add a number of standard deviations worth of bid (ask) volume at or near the level two price, but less than the level one price to push the price up (down).

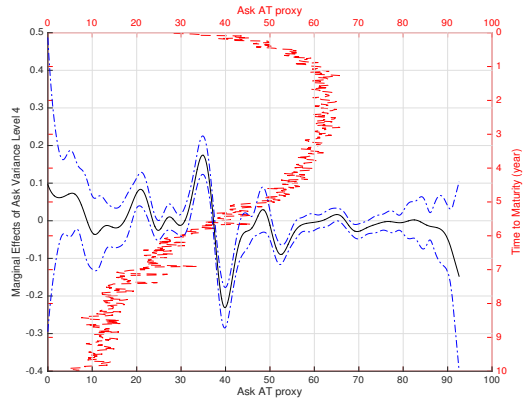
I see that the volume of messages in the limit order book at level 2 (not to be conflated with the level 2 wavelet decomposition) is by far the highest, approaching double that of the level 1 messages at or around two years from maturity. However, the peaks for the bids and the asks are not strictly concurrent and I suggest that it is this aesthetic choice in which the side of the order-book in which the algorithms with operate is a key driver of the asymmetries between the bid and ask AT proxies impact on the various market characteristics and in particular that level one variance and covariance ratios and effect that is strongly suggested in Figure 4.3.

As previously indicated, the variance ratios at medium frequencies have significant impacts at execution risk. Figure 4.4 Subplot (a) to (f) reveal the marginal effects of algorithmic trading on variance and covariance ratios at level 4 with the panel dataset. Subplots (a) and (b) indicate that AT has impacts on the asks side variance ratio depending on the proportion of algorithmic trading active in the market. As ask side AT proxy approaches around 32% and 35%, there is a substantial increase in the marginal effects for the ask variance ratio level 4, which indicates that algorithmic traders increase the ask-side volatility and

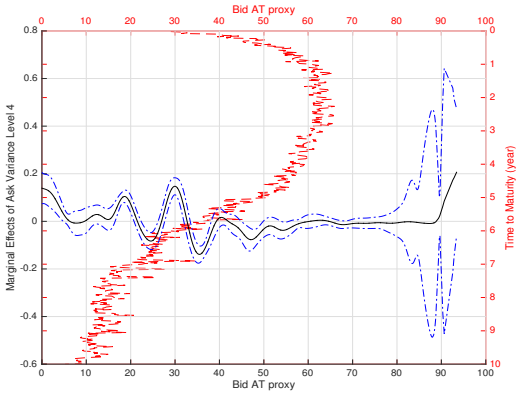
provide substantial microstructure noises into the markets, recall that by level four, the variance scaling is a comparison to a degree of 1/16 of trading day, hence the local volatility is substantially scaled at this point.

Once the proportion of ask side AT increases further around 36% to 40%, the marginal effects for the ask variance ratio decrease substantially. This could be caused by the ATs saturate the market and their orders start to cancel each other out. Hence, the liquidity providing from AT will drop off and the spreads variance will shrink. As AT proxy gets higher, they will fully saturate the market, and eventually no extra effects on the variance ratios. Subplots (c) and (d) present the effects of AT on the bids side variance ratios at level 4. Compared to the ask variance ratios, the critical points of bids variance happen earlier, around 20% to 40%.

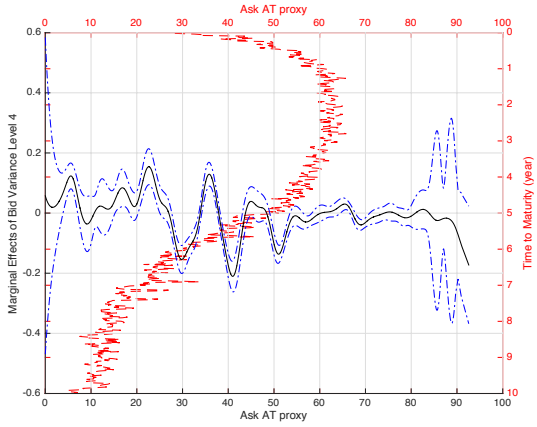
In Subplots (e) and (f), the marginal effects of AT proxy on bid–ask covariance ratio level 4 also illustrate the decoupling of the bid and ask at medium frequencies. The ask side AT proxy has no impact on the bid-ask covariance ratio at wavelet level 4 when the proportion of ask AT is less than 30% and after it increases further than 48% approximately. This tells a similar story with bid–ask covariance ratio at level 1. Within a 30% to 33% proportion the impact of ask side AT is positive as the proportion of AT decreases the decoupling of the bid and ask. The proportion between 33%and 40% AT substantially decreases the covariance ratio, therefore the fraction of AT leads to a significant decoupling of the bid-ask spreads. For the impacts of bid side AT proxy on bid–ask covariance ratio at level 4, the decoupling happens much at a lower proportion of AT, around 13% to 27% (see Figure 4.4 Subplot (f)). Broadly speaking, Figure 4.4 reveals that the proportion of AT activity in the market has saturation points, and decouples bid and ask movements.



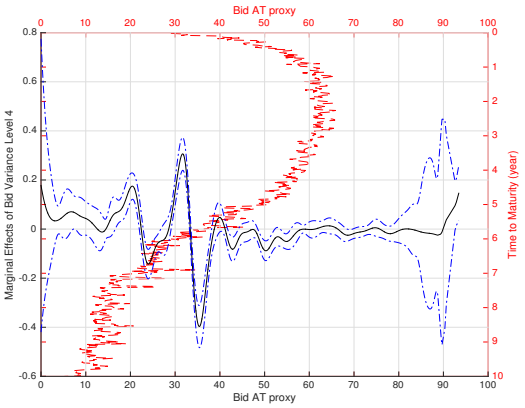
(a) Ask Variance Level 4 - Panel



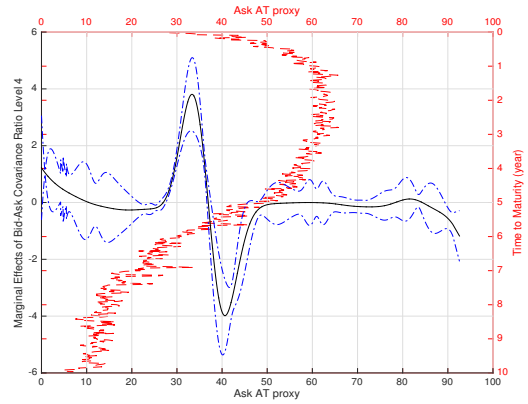
(b) Ask Variance Level 4 - Panel



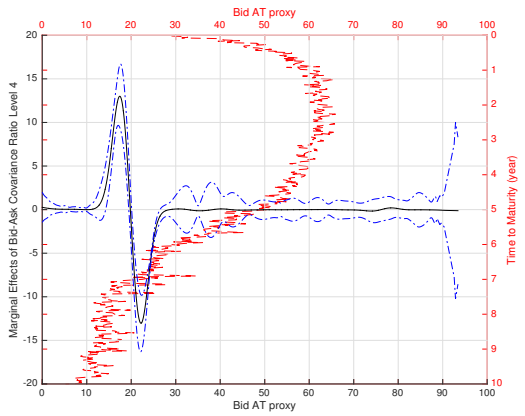
(c) Bid Variance Level 4 - Panel



(d) Bid Variance Level 4 - Panel



(e) Bid Ask Covariance Level 4 - Panel



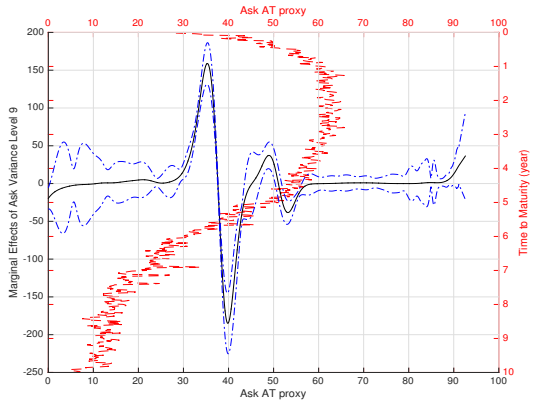
(f) Bid Ask Covariance Level 4 - Panel

Figure 4.4: Marginal Effects of Variance and Covariance at Timescale Level 4 Thresholded at 25 ms

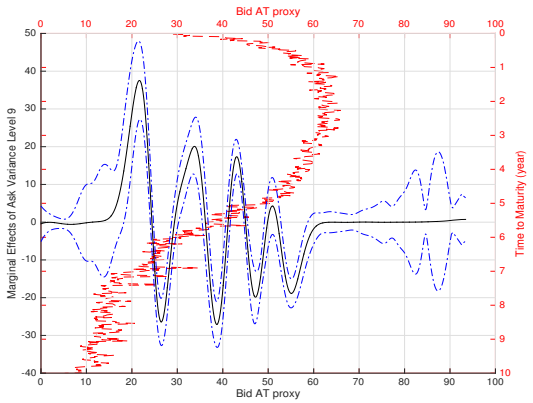
Notes: The red dashed line represents the evolution of daily AT proxy thresholded at 25 ms for order book level 2 in function of the average time to maturity. The black continuous line depicts the marginal effects of AT on variance ratio in Subplots (a) to (d) and, on covariance ratio in Subplots (e) and (f), both at timescale 4 with the panel dataset. The blue dashed lines are the 95% lower and upper confidence bounds of marginal effects.

For longer time frequencies, Figure 4.5 illustrates the effects of algorithmic trading variance ratio in Subplots (a) to (d) and bid–ask covariance ratio in Subplots (e) and (f) at the longest timestamp level 9 with the panel dataset. When the fraction of ask side AT proxy is lower than 30%, there is a limited influence of AT on ask variance ratio at level 9. Within a proportion of 30% to 35% ask side AT, there is a substantial increase in the marginal effects for the ask variance ratio. Once the ask AT proxy increases around 36% to 40%, the marginal effects of AT on ask variance ratio substantially decrease. When the fraction of AT proxy is over 44%, the marginal effects shift around and the algorithmic trading activity in the market has limited influence on the ask variance ratio at level 9. For the impacts coming from the bid side AT proportion, there is no significant effects on ask variance ratio when the fraction of bid side AT is below 16% and above 60%. However, the bid side AT proxy has positive effects on the ask variance ratio within a proportion of 17% to 22%, 26%to 34% and 39% to 43%. In these periods, algorithmic traders play a liquidity supplier role in the market, and result in the increasing both market liquidity and variance ratio.

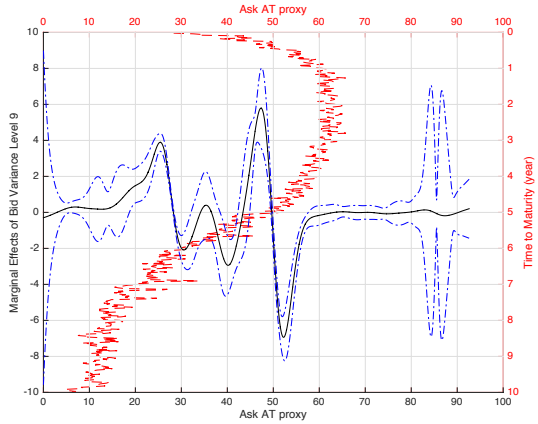
Following each marginal effects increasing period, there is also a drop down period for the ask variance ratio in the rest proportion within 44% and 55%. The bid side AT proxy brings negative impacts on the ask variance ratio and results in dramatically decreases. The reason could be that the market reaches its saturation limits and orders from bid side ATs start to cancel each other out, therefore the liquidity contracts and the ask variance decreases. The effects from AT proportion on the bid variance ratio at level 9 display more fluctuations and less extreme than the ask variance ratio, and become stable above the proportion of 60% ATs. Subplots (e) and (f) describe that the fraction of AT has the ability to decouple between bids and asks of the book. The evidence indicates that ask side AT has limited impacts on the bid–ask covariance ratio at time scale level



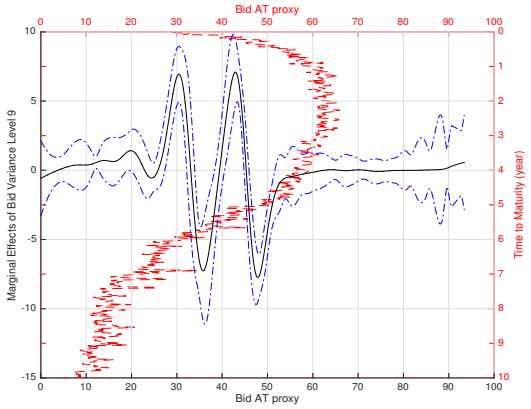
(a) Ask Variance Level 9 - Panel



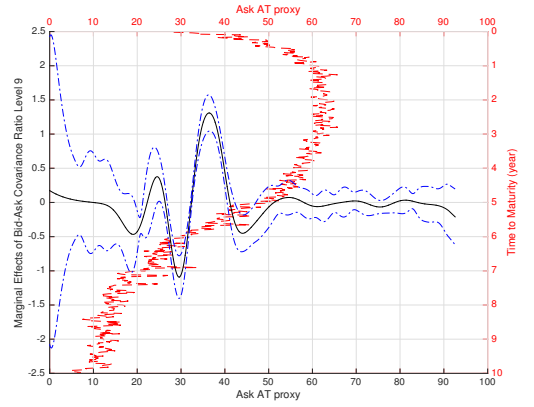
(b) Ask Variance Level 9 - Panel



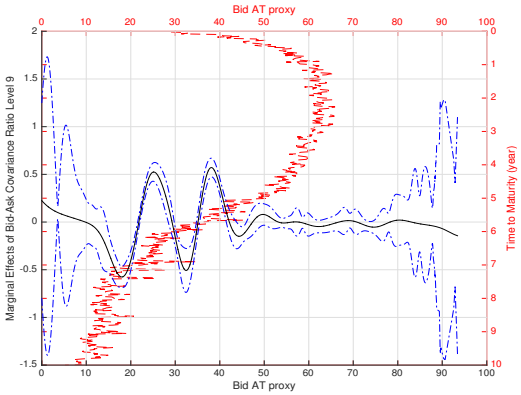
(c) Bid Variance Level 9 - Panel



(d) Bid Variance Level 9 - Panel



(e) Bid Ask Covariance Level 9 - Panel



(f) Bid Ask Covariance Level 9 - Panel

Figure 4.5: Marginal Effects of Variance and Covariance at Timescale Level 9 Thresholded at 25 ms

Notes: The red dashed line represents the evolution of daily AT proxy thresholded at 25 ms for order book level 2 in function of the average time to maturity. The black continuous line depicts the marginal effects of AT on variance ratio in Subplots (a) to (d) and, on covariance ratio in Subplots (e) and (f), both at timescale 9 with the panel dataset. The blue dashed lines are the 95% lower and upper confidence bounds of marginal effects.

9 before the proportion of AT below 14% and above 52% approximately. When the fraction of ask side AT is between 14% - 19%, 25% - 29% and 37% - 43%, the impact of AT on bid–ask covariance ratio is negative. This means the proportion of AT results in the bid–ask spreads decoupling. When the fraction of ask side AT is within 20% - 25% and 30% - 36%, the AT substantially increases the bid–ask covariance ratio, which indicates that the proportion of AT decreases the decoupling degrees of the bid–ask spreads. For the impacts from bid side AT activities, when the bid AT proxy increases to 18% - 26% and 33% - 38%, the marginal effect on bid–ask covariance ratio is substantially increased. Once the proportion of AT is between 27% - 32% and 38% - 42%, the bid–ask covariance ratio dips and results in the increasing degrees of bid–ask decoupling.

4.3.2 Marginal Effects of AT Proxy Thresholded at 200ms

In this section, I discuss the marginal effects of algorithmic trading proxy thresholded at 200ms on the diverse market quality indicators. Table 4.2 records the results of my semi-parametric regressions with AT proxy thresholded at 200ms. The results is similar with the one thresholded at 25ms. Table 4.2 illustrates the case that, when the AT proxy is regressed against the standard measures of liquidity the coefficients are generally positively and significant; only when regressed against the lagged effective spread is the AT proxy not significant for the volume weighted average dataset. For the remaining cases, the coefficients attached to AT range from 0.065 to 0.917 for the ask side and from 0.073 to 0.920 for the bid side and, are all significant at a 99% confidence level. When compared, side-by-side, the coefficients on the AT proxy are slightly higher for the bid side when using the volume weighted dataset. This effect is not observed in the panel dataset where there is almost no difference between the explanatory

power of AT over execution risk between the bid and the ask sides.

Figure 4.6 to 4.10 presents parts of the estimated non-parametric function with the algorithmic trading thresholded at 200ms. In Figure 4.6, the black line represents the marginal effects of algorithmic trading on the bid-ask spreads with AT proxy thresholded at 200 ms. The marginal effects of algorithmic trading on the bid-ask spreads have been measured by two set of data, one is the volume weighted average dataset in Subplots (a) and (b), the other is the panel dataset in Subplots (c) and (d). The red dashed line shows the proportion of daily AT proxy thresholded at 200ms for the order book level 2.

Figure 4.6 Subplots (a) to (d) indicate that the increase of high-speeding algorithmic traders has effects on the bid-ask spreads. In Subplot (b) the marginal effects of the bid-side algorithmic trading fraction on the bid-ask spreads reach at the peak when the bid-side AT proxy at 26%. Once the AT fraction across 26%, the marginal effects on the bid-ask spreads suddenly drop to the bottom around 33% proportion. This could be the high-frequency algorithmic traders provide liquidity to the market, decrease the transaction costs and narrow bid-ask spread. Once the bid-side AT fraction passes 33%, the impact of bid side AT proxy on the bid-ask spreads spikes until the AT fraction reaching around 37%. The increasing of algorithmic traders could create more informed traders existed in the market, which could lead to higher transaction costs, lower market liquidity and widen bid-ask spreads. It is worth to note that there is another critical point when the bid-side algorithmic traders fraction is around 66%. Furthermore, Subplot (c) and (d) capture the non-parametric results from the panel dataset. For both ask-side and bid-side, the increase of high frequency algorithmic traders has the influence on the bid-ask spreads when the AT proxy is below 50%.

In Figure 4.7 the black line represents the marginal effects of algorithmic trading

Table 4.2: The Semi-parametric Regressions Results with AT Proxy Thresholded at 200ms

		Liqudity Measures								Variance Ratios								
		$\tilde{\mathscr{S}}^{Q}$	$\tilde{\mathscr{S}}^{Q^{1/2}}$	$\tilde{\mathscr{S}}^{D}$	$\tilde{\mathscr{S}}^{E}$	$\tilde{\mathscr{S}}^{R,5}$	$\tilde{\mathscr{S}}^{R,30}$	$\tilde{\mathscr{S}}^{AS,5}$	$\tilde{\mathscr{S}}^{AS,30}$	$\tilde{\mathscr{R}}_{A1}$	$\tilde{\mathscr{R}}_{B1}$	$\tilde{\mathscr{R}}_{AB1}$	$\tilde{\mathscr{R}}_{A4}$	$\tilde{\mathscr{R}}_{B4}$	$\tilde{\mathscr{R}}_{AB4}$	$\tilde{\mathscr{R}}_{A9}$	$\tilde{\mathscr{R}}_{B9}$	$\tilde{\mathscr{R}}_{AB9}$
Volume weighted average																		
	$L(Y_{i,t})$	0.358	0.358	0.917	-0.005	0.065	0.571	0.132	0.623	-0.009	-0.002	-0.001	-0.001	0.007	-0.001	-0.006	-0.004	0.003
		(19.65)**	(19.65)**	(131.57)**	-0.30	(3.18)**	(33.92)**	(6.42)**	(38.31)**	-0.41	-0.10	-0.06	-0.04	-0.36	-0.07	-0.27	-0.21	-0.16
Ask	$W_{i,t}$	-0.078	-0.039	0.016	0.006	0.010	0.012	-0.003	-0.005	0.007	0.010	0.340	-0.015	0.009	-1.076	4.141	0.755	0.056
AT Proxy		(3.87)**	(3.87)**	-0.95	(2.06)*	-0.73	-0.97	-0.21	-0.44	(2.36)*	(2.51)*	(2.19)*	-0.54	-0.38	(2.16)*	-1.18	(2.31)*	-1.03
	R^2	0.15	0.15	0.88	0.00	0.00	0.33	0.02	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339
	$L(Y_{i,t})$	0.373	0.373	0.920	0.104	0.073	0.578	0.133	0.624	-0.013	-0.006	-0.010	-0.003	0.006	-0.011	-0.006	-0.003	0.004
		(20.02)**	(20.02)**	(133.81)**	(5.17)**	(3.58)**	(34.23)**	(6.50)**	(38.47)**	-0.64	-0.29	-0.48	-0.15	-0.28	-0.52	-0.27	-0.13	-0.17
Bid	$W_{i,t}$	-0.067	-0.034	-0.025	-0.001	-0.006	0.001	0.004	0.003	0.011	0.018	0.309	0.006	0.070	-0.947	2.927	1.014	0.057
AT Proxy		(3.48)**	(3.48)**	-1.62	-0.40	-0.51	-0.11	-0.35	-0.23	(3.99)**	(4.90)**	(2.14)*	-0.25	(3.12)**	(2.05)*	-0.90	(3.36)**	-1.13
	R^2	0.15	0.15	0.89	0.01	0.01	0.33	0.02	0.39	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339	2,339
Panel																		
	$L(Y_{i,t})$	0.735	0.735	0.881	0.601	0.074	0.590	0.131	0.613	0.004	0.002	0.000	0.047	0.123	0.000	0.001	0.003	0.006
		(205.94)**	(205.94)**	(377.78)**	(141.74)**	(13.95)**	(137.76)**	(24.93)**	(146.19)**	-0.82	-0.45	-0.03	(8.94)**	(23.37)**	-0.02	-0.24	-0.56	-1.08
Ask	$W_{i,t}$	0.137	0.068	-0.397	-0.008	0.029	0.026	-0.045	-0.033	0.005	0.002	-0.291	0.002	-0.004	0.886	-6.265	-0.836	0.080
AT Proxy		(14.01)**	(14.01)**	(32.22)**	(7.98)**	(2.31)*	(2.29)*	(3.41)**	(2.86)**	(3.01)**	-0.98	(3.25)**	-0.35	-0.80	(3.09)**	(2.01)*	(2.43)*	-1.33
	R^2	0.55	0.55	0.83	0.37	0.01	0.35	0.02	0.38	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
	N	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491
	$L(Y_{i,t})$	0.736	0.736	0.880	0.601	0.074	0.590	0.131	0.613	0.005	0.003	0.000	0.047	0.122	0.000	0.001	0.003	0.006
		(206.19)**	(206.19)**	(377.96)**	(141.73)**	(13.98)**	(137.79)**	(24.97)**	(146.23)**	-0.91	-0.52	0.00	(8.81)**	(23.16)**	-0.01	-0.23	-0.49	-1.12
Bid	$W_{i,t}$	0.123	0.062	-0.406	-0.008	0.020	0.010	-0.036	-0.017	0.004	0.003	0.076	-0.005	0.004	-0.279	-2.968	-0.761	0.082
AT Proxy		(12.62)**	(12.62)**	(32.96)**	(7.81)**	-1.61	-0.88	(2.73)**	-1.45	(2.62)**	-1.22	-0.85	-0.84	-0.87	-0.97	-0.95	(2.21)*	-1.36
	R^2	0.55	0.55	0.83	0.37	0.01	0.35	0.02	0.38	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	N	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491	35,491

* p <0.05; ** p <0.01

Notes: This table illustrates coefficients estimate results from the semi-parametric regression: $Y_{i,t} = X_{i,t}\beta + \mathscr{G}(Z_{i,t}) + \varepsilon_{i,t}$, where $Y_{i,t}$ is the measure of execution risk/market quality for the i contract type for day t, $X_{i,t}$ are my $k = 2$ linear regressors, $Z_{i,t}$ is my $p = 1$ algorithmic trading proxy thresholded at 200ms, and $W_{i,t}$ is a set of instrumental variables, where $\mathbb{E}[\varepsilon_{i,t}|W_{i,t}] = 0$. Flattening the data set to an index $i \in \{1, \ldots, N\}$, where N is the total number of day-contracts and the index i represents the set of N tuples of (i, t) contract-days. In this table, the dependent variables measuring the execution risk $Y_{i,t}$ are reported into two groups: liquidity measures and variance/covariance ratios. Liquidity measurements contain the bid-ask spreads ($\tilde{\mathscr{S}}^{Q}$), quoted half spreads ($\tilde{\mathscr{S}}^{Q^{1/2}}$), quoted depth ($\tilde{\mathscr{S}}^{D}$), effective spreads ($\tilde{\mathscr{S}}^{E}$), realized spreads 5 minutes ($\tilde{\mathscr{S}}^{R,5}$), realized spreads 30 minutes ($\tilde{\mathscr{S}}^{R,30}$), adverse selection 5 minutes ($\tilde{\mathscr{S}}^{AS,5}$) and adverse selection 30 minutes ($\tilde{\mathscr{S}}^{AS,30}$). Variance/covariance ratios include ask variance ratios ($\tilde{\mathscr{R}}_{A}$), bid variance ratios ($\tilde{\mathscr{R}}_{B}$) and bid-ask covariance ratios ($\tilde{\mathscr{R}}_{AB}$) at level 1, level 4 and level 9. The lagged one dependent variable is denoted as $L(Y_{i,t})$. Instrumental variables are adopted to deal with the endogeneity problems, named as $W_{i,t}$. The volume weighted average dataset uses algorithmic trading proxy instrumented on the log of total messages, and the panel dataset adopts algorithmic trading proxy instrumented on both the log of total messages and the log of time to maturity.

on adverse selection (5 mins) in Subplots (a) and (b) and, on adverse selection (30 mins) in Subplots (c) and (d), both with the volume weighted average dataset. The red dashed line shows the proportion of daily AT proxy for the order book level 2. Figure 4.7 indicates that the increasing AT activities have effects on both adverse selection within both 5 minutes and 30 minutes. Subplots (a) and (b) suggest that AT has limited effects on the adverse selection (5 mins) when the AT proxy below 60%. Once the AT proxy increases over 60%, the effect of AT on adverse selection (5 mins) increases dramatically to the peak. This indicates that ATs increase trading costs, decrease the market liquidity, and damage the price efficiency with more informative quotes in the market. At this stage, the increasing algorithmic traders act as informed traders to submit and cancel quotes orders. The degree of asymmetric information in the market reaches a higher level with increased the transaction cost to execute orders. However, the marginal effects on adverse selection (5min) suddenly drops, as the ask-side AT proxy increases to about 73% - 78%. In this period, ATs provide liquidity to the market and increase price efficiency with less informative quotes.

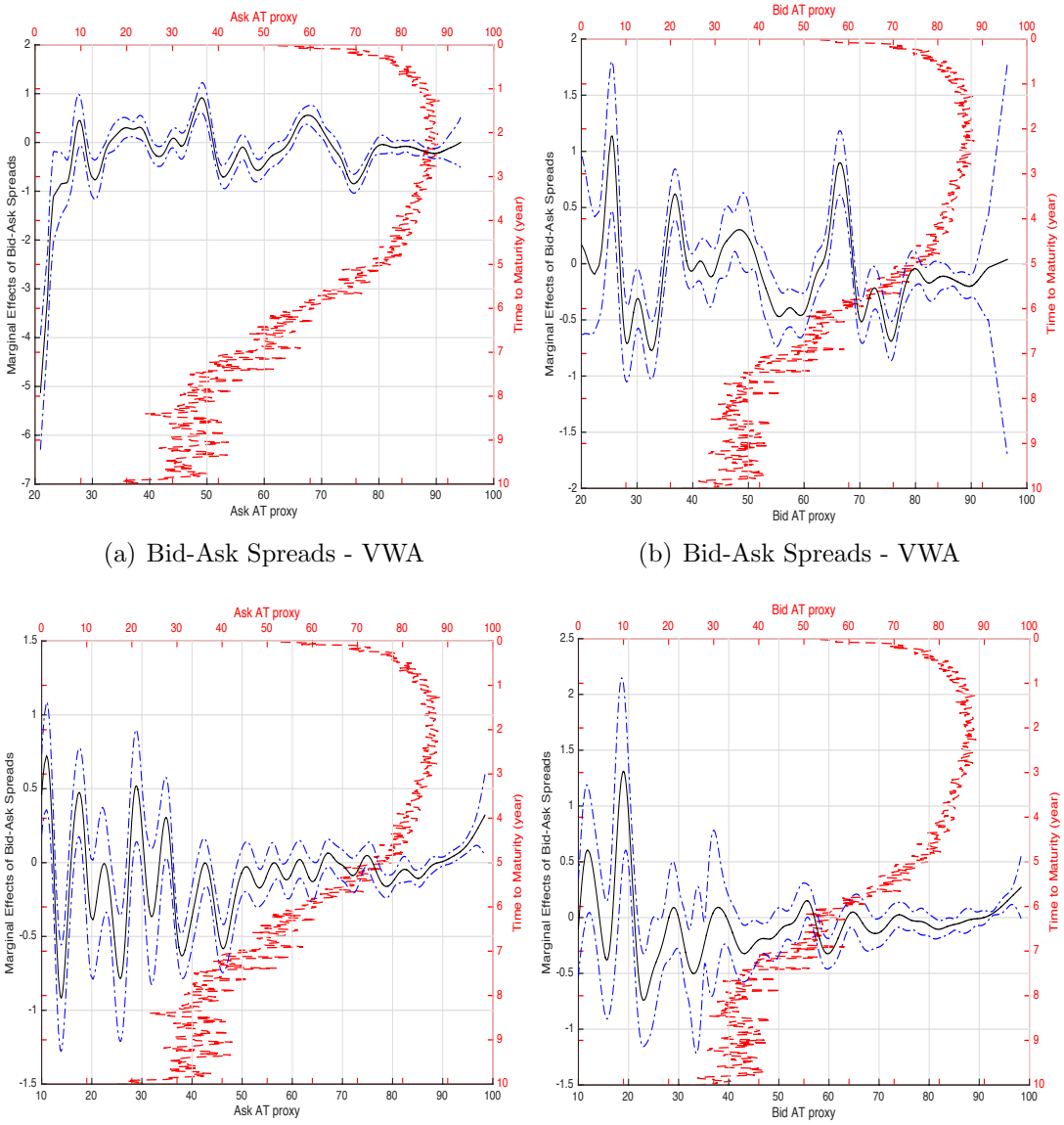
In Subplots (c) and (d), the effects of AT proxy on adverse selection (30 mins) show the similar patterns. There are no distinct effects of bid side AT proxy on the adverse selection (30 mins) when the proportion of algorithmic traders is below around 60% (see Figure 4.7 Subplot (d)). Within 60% to 67% proportion, the marginal effects are positive. Once the AT fraction passes 67%, the impact of bid side AT proxy on adverse selection (30 mins) starts to drop until the AT fraction reaching around 77%.

Overall, the proportion of algorithmic trading significantly impacts on the adverse selection (5 mins and 30 mins) and market liquidity. AT shows limited impacts when the AT proxy is relative low. As the increase of AT proportion,

the algorithmic traders act as liquidity demanders, which ATs act as informed traders and its impacts on adverse selection increase. Once the AT reaches a certain critical point, the algorithmic traders act as liquidity providers, which leads to the shrinkage of asymmetry information and the increasing market liquidity. When the most traders in the market are ATs, the marginal impacts effectively drop to zero.

In Figure 4.8 the dashed red line represents the evolution of daily AT proxy for order book level 2 in function of the average time to maturity. The blue continuous line depicts the marginal effects of AT on variance ratio in Subplots (a) to (d) and, on covariance ratio in Subplots (e) and (f), both at timescale 1 with the panel dataset. Subplots (a) and (b) suggest that AT effects on both the variance and the covariance ratios depend on the proportion of algorithmic trading active in the market. When this proportion is lower than approximately 40%, the evidence suggests that AT has limited effect on the variance ratio of the spreads. Once this proportion of AT is higher than 20%, there is an increase in the marginal effect on the variance ratio revealing AT increase the high-frequency microstructure noises and volatilities in the market. At the same time, higher variance ratios are symptomatic of less informative quotes and, therefore, higher execution risk as the fundamental price of the contracts is more difficult to observe in the market. However, as the proportion of AT increases further (to around reaches 48 - 50%) the effect on the variance ratio drops dramatically most likely because the high proportion of AT results in their orders cancelling each other, shrinking liquidity and, in turn, the spreads variance.

The marginal effects for the bid variance ratio level 1 in Subplots (c) and (d) present the similar pattern compared with the ask variance ratio. When the ask side AT proxy is below 42%, the impact of algorithmic trading on bid variance



(a) Bid-Ask Spreads - VWA

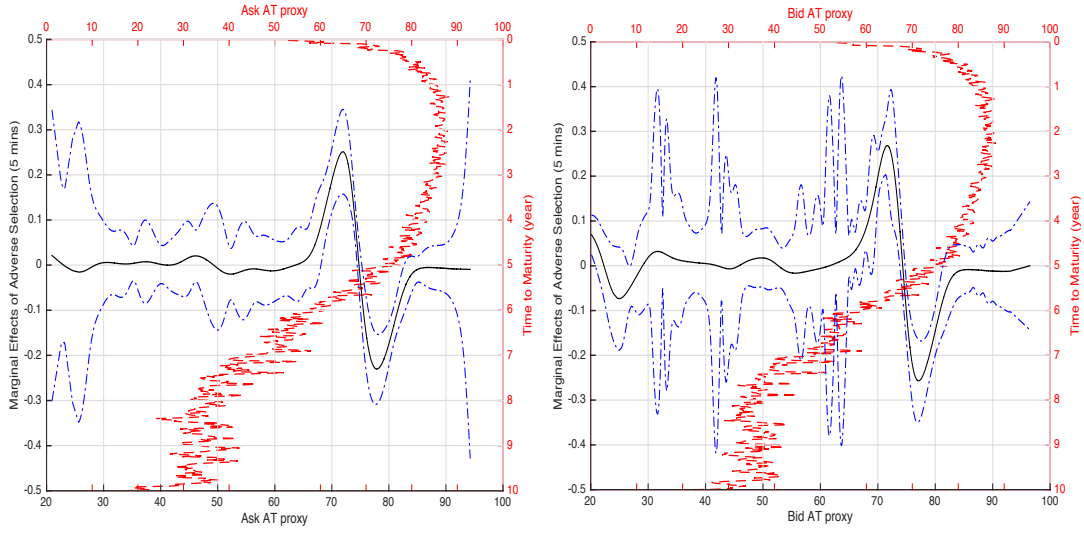
(b) Bid-Ask Spreads - VWA

(c) Bid-Ask Spreads - Panel

(d) Bid-Ask Spreads - Panel

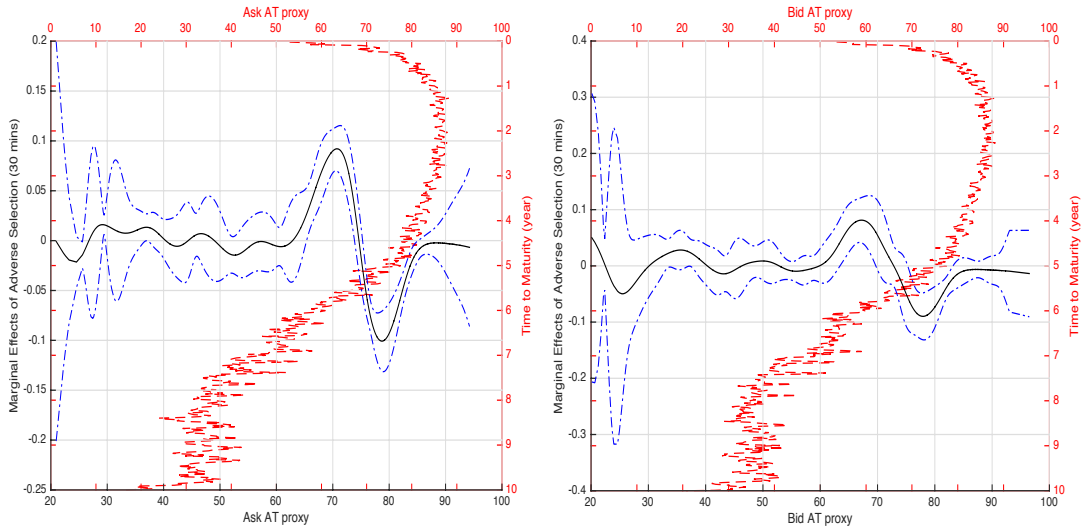
Figure 4.6: Marginal Effects of Bid-Ask Spreads Thresholded at 200 ms

Notes: The red dashed line represents the evolution of daily AT proxy thresholded at 200 ms for order book level 2 in function of the average time to maturity. The black continuous line depicts the marginal effects of AT on the bid-ask spreads in Subplots (a) to (b) with volume weighted average dataset and Subplots (c) and (d) with the panel dataset. The blue dashed lines are the 95% lower and upper confidence bounds of bid-ask spreads marginal effects.



(a) Adverse Selection (5 mins) - VWA

(b) Adverse Selection (5 mins) - VWA



(c) Adverse Selection (30 mins) - VWA

(d) Adverse Selection (30 mins) - VWA

Figure 4.7: Marginal Effects of Adverse Selection (5 mins and 30 mins) Thresholded at 200 ms

Notes: The red dashed line represents the evolution of daily AT proxy thresholded at 200 ms for order book level 2 in function of the average time to maturity. The black continuous line depicts the marginal effects of AT on adverse selection (5 mins) in Subplots (a) to (b) and, on adverse selection (30 mins) in Subplots (c) and (d), both with the volume weighted average dataset. The blue dashed lines are the 95% lower and upper confidence bounds of adverse selection marginal effects.

ratio is relatively weak. When the proportion of ATs is around 43% to 47%, its impacts on bid variance suddenly spike to the top. Once the fraction of ATs is higher than 47%, the marginal effects drop dramatically. Then the marginal effects on the bid variance ratio level 1 increase, when the ask side AT proxy passes around 52%. Eventually there are no much effects on the bid variance ratio at level 1, when the ask side AT proxy is higher than 56%. For the impact of the bid side AT proxy on the bid variance ratio at level 1, the pattern is similar to the ask side; however, the point of inflection is lower than for the ask side, at around 33% - 43% of messages being ascribed to be algorithmic in origin.

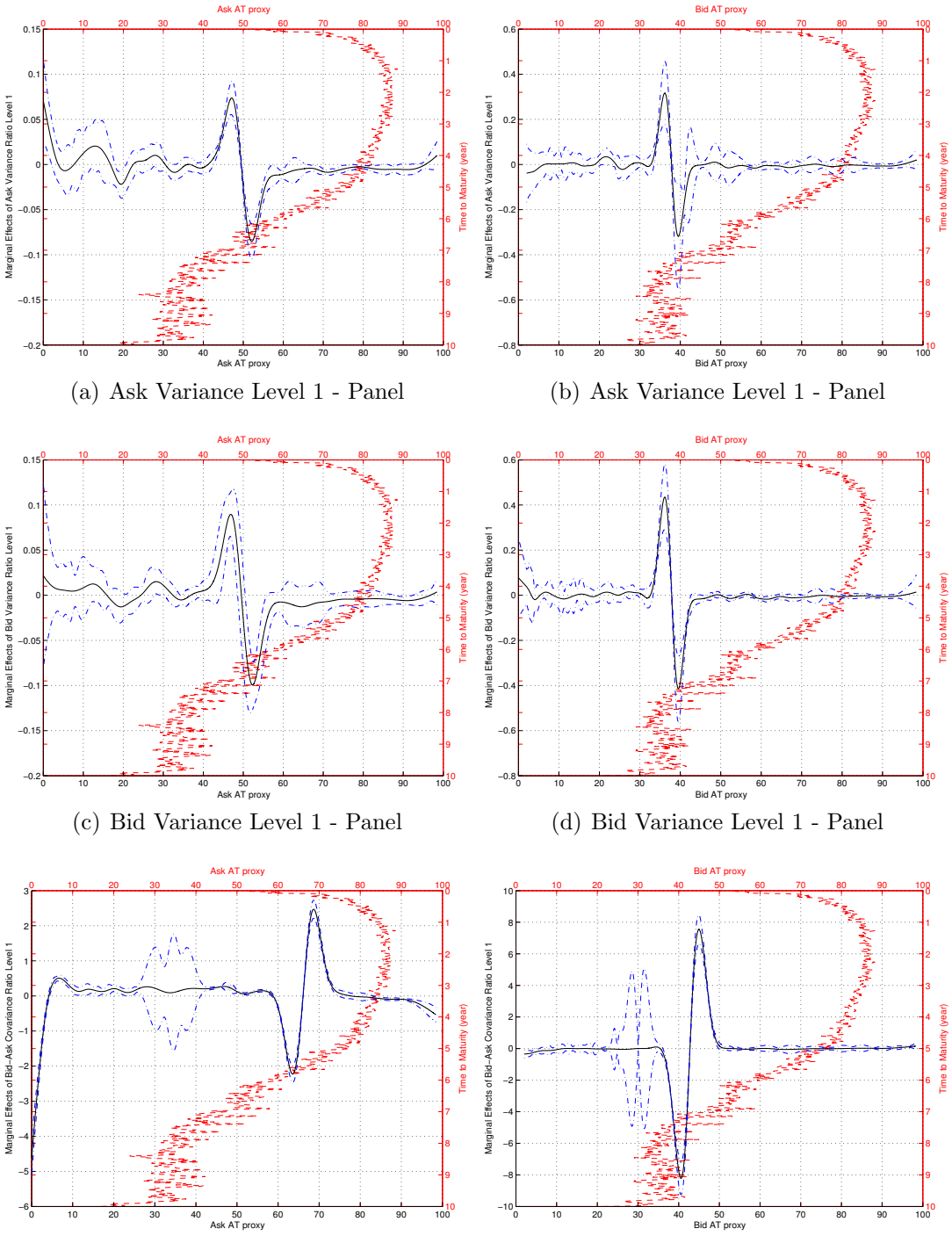
The evidence suggests that once the AT proxy reaches the critical point 43%, variance ratios will stabilize. I conjecture that this is because once the majority of quoted volume belongs to algorithmic traders using the similar algorithms the arms race has saturated the market as each sends out the similar signals. I define this critical point as AT market saturation, I believe that it is this population measurement that explains the differing results from prior studies in the literature that have only had access to short snippets of the order-flow over say a single month or in some cases only a few days.

Subplots (e) and (f) tell a similar story but in terms of decoupling of bids and asks. As in the previous case, the effect of AT on the covariance ratio at time scale 1 are virtually non-existent before the proportion of AT reaches around 60% and after it increases further than 72-75% approximately. Within a 60% to 65% proportion the impact of AT is negative as the proportion of AT results in a significant decoupling of bid - ask spreads. Then, up to a proportion of 75% AT substantially increases the covariance ratio, therefore reducing the decoupling; one of key interest here is the shape of the transmission function for the covariance versus the variance, notice that in keeping with the saturation argu-

ment the deterioration in market quality (the reduction in covariation evidencing a decoupling) occurs at lower fractions of high speed messages. However, the points of inflection for the bids versus the asks are markedly different and indeed these points differ from the inflection points for the variances. The reduction in covariation of the level 1 bid-ask spread is impacted at a substantially lower fraction of high speed AT messages than for the bid AT proxy. The separation of the volatility and correlation effects at level one is quite fascinating and indicates that the mix strategies employed by HFTs obviously have markedly differentiated impacts.

As previously indicated, the variance ratios at medium frequencies have significant impacts at execution risk. Figure 4.9 reveals the marginal effects of algorithmic trading on variance and covariance ratios at level 4 with the panel dataset. Subplots (a) and (b) indicate that AT has impacts on the asks side variance ratio depending on the proportion of algorithmic trading active in the market. As ask side AT proxy approaches around 57% and 62%, there is a substantial increase in the marginal effects for the ask variance ratio level 4, which indicates that algorithmic traders increase the ask-side volatility and provide substantial microstructure noises into the markets. The higher variance ratios indicate the more opportunities for a sophisticated trader to arbitrage.

Once the proportion of ask side AT increases further around 63% to 76%, the marginal effects for the ask variance ratio decrease substantially. This could be caused by the ATs saturate the market and their orders start to cancel each other out. Hence, the liquidity providing from AT will drop off and the spreads variance will shrink. As AT proxy gets higher, they will fully saturate the market, and eventually no extra effects on the variance ratios. Subplots (a) and (b) present the effects of AT on the bids side variance ratios at level 4. Compared to the ask



(a) Ask Variance Level 1 - Panel

(b) Ask Variance Level 1 - Panel

(c) Bid Variance Level 1 - Panel

(d) Bid Variance Level 1 - Panel

(e) Bid Ask Covariance Level 1 - Panel

(f) Bid Ask Covariance Level 1 - Panel

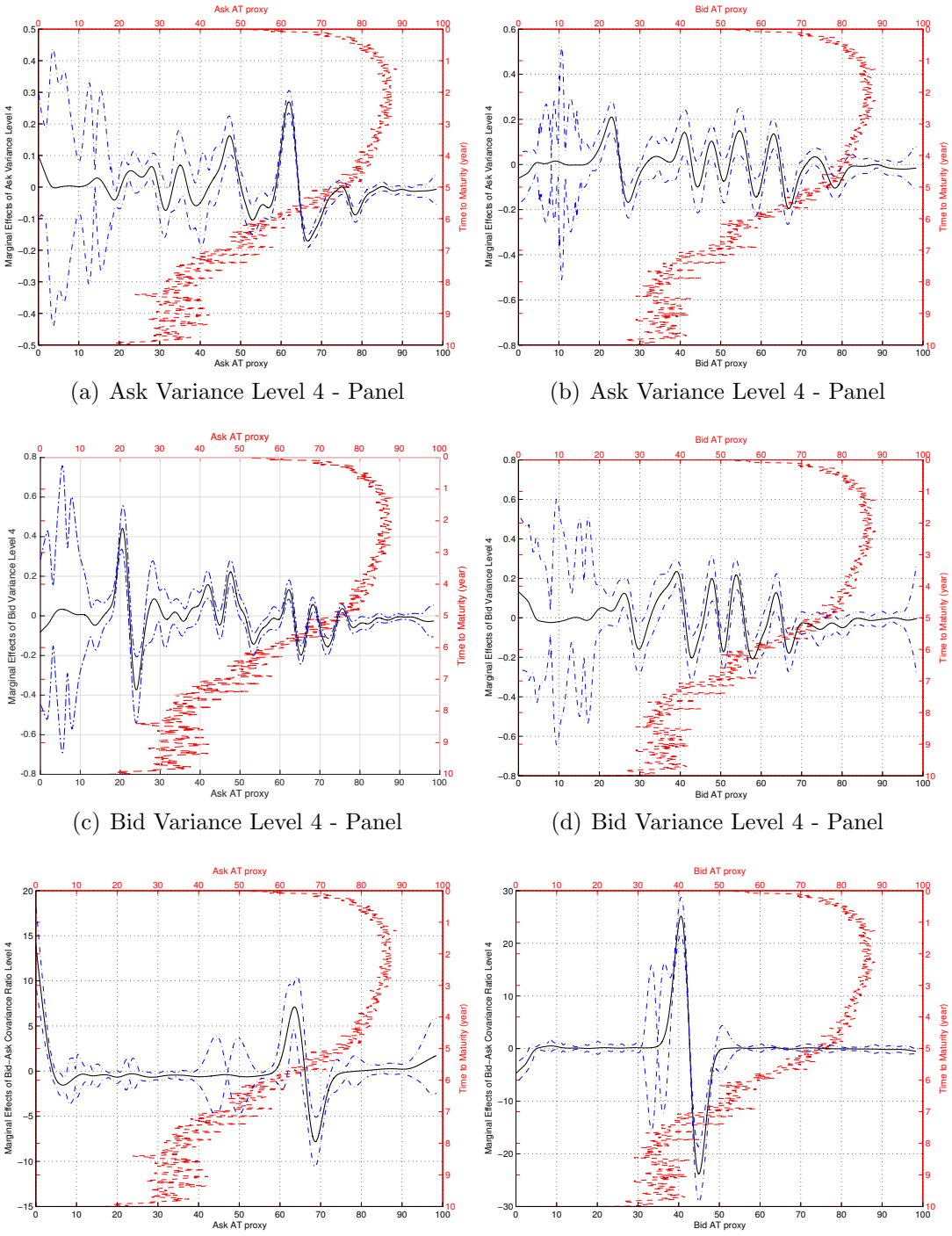
Figure 4.8: Marginal Effects of Variance and Covariance at Timescale Level 1 Thresholded at 200 ms

Notes: The red dashed line represents the evolution of daily AT proxy thresholded at 200 ms for order book level 2 in function of the average time to maturity. The black continuous line depicts the marginal effects of AT on variance ratio in Subplots (a) to (d) and, on covariance ratio in Subplots (e) and (f), both at timescale 1 with the panel dataset. The blue dashed lines are the 95% lower and upper confidence bounds of marginal effects.

variance ratios, the critical points of bids variance happen much earlier, around 20% to 25%.

In Subplots (e) and (f), the marginal effects of AT proxy on bid–ask covariance ratio level 4 also illustrate the decoupling of the bid and ask at medium frequencies. The ask side AT proxy has no impact on the bid-ask covariance ratio at wavelet level 4 when the proportion of ask AT is less than 60% and after it increases further than 72% approximately. This tells a similar story with bid–ask covariance ratio at level 1. Within a 60% to 64% proportion the impact of ask side AT is positive as the proportion of AT decreases the decoupling of the bid and ask. The proportion between 65%and 72% AT substantially decreases the covariance ratio, therefore the fraction of AT leads to a significant decoupling of the bid-ask spreads. For the impacts of bid side AT proxy on bid–ask covariance ratio at level 4, the decoupling happens much at a lower proportion of AT, around 41% to 45% (see Figure 4.9 Subplot (f)). Broadly speaking, Figure 4.9 reveals that the proportion of AT active in the market has saturation points, and decouples bid and ask movements.

For longer time frequencies, Figure 4.10 illustrates the effects of algorithmic trading variance ratio in Subplots (a) to (d) and bid–ask covariance ratio in Subplots (e) and (f) at the longest timestamp level 9 with the panel dataset. When the fraction of ask side AT proxy is lower than 55%, there is a limited influence of AT on ask variance ratio at level 9. Within a proportion of 55% to 61% ask side AT, there is a substantial increase in the marginal effects for the ask variance ratio. Once the ask AT proxy increases around 62% to 66%, the marginal effects of AT on ask variance ratio substantially decrease. When the fraction of AT proxy is over 66%, the marginal effects shift around and the algorithmic trading activity in the market has limited influence on the ask variance ratio at level 9. For the



(a) Ask Variance Level 4 - Panel (b) Ask Variance Level 4 - Panel

(c) Bid Variance Level 4 - Panel (d) Bid Variance Level 4 - Panel

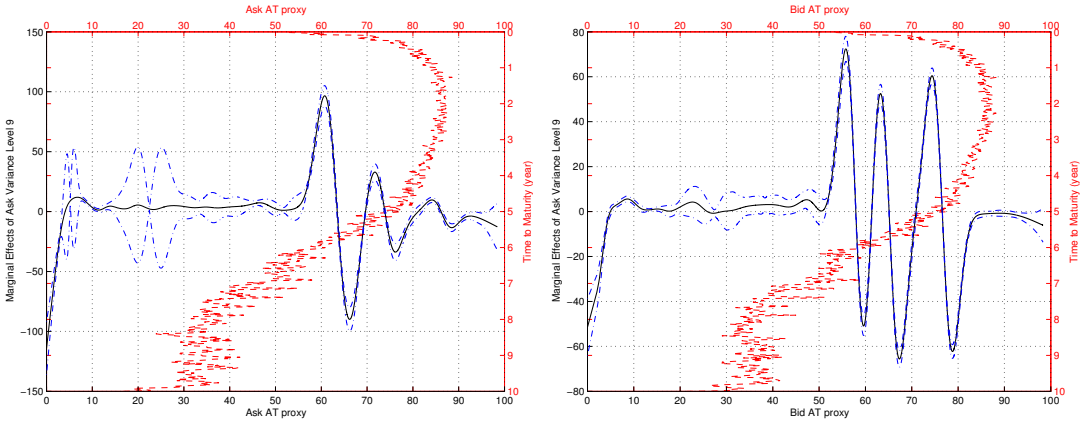
(e) Bid Ask Covariance Level 4 - Panel (f) Bid Ask Covariance Level 4 - Panel

Figure 4.9: Marginal Effects of Variance and Covariance at Timescale Level 4 Thresholded at 200 ms

Notes: The red dashed line represents the evolution of daily AT proxy thresholded at 200 ms for order book level 2 in function of the average time to maturity. The black continuous line depicts the marginal effects of AT on variance ratio in Subplots (a) to (d) and, on covariance ratio in Subplots (e) and (f), both at timescale 4 with the panel dataset. The blue dashed lines are the 95% lower and upper confidence bounds of marginal effects.

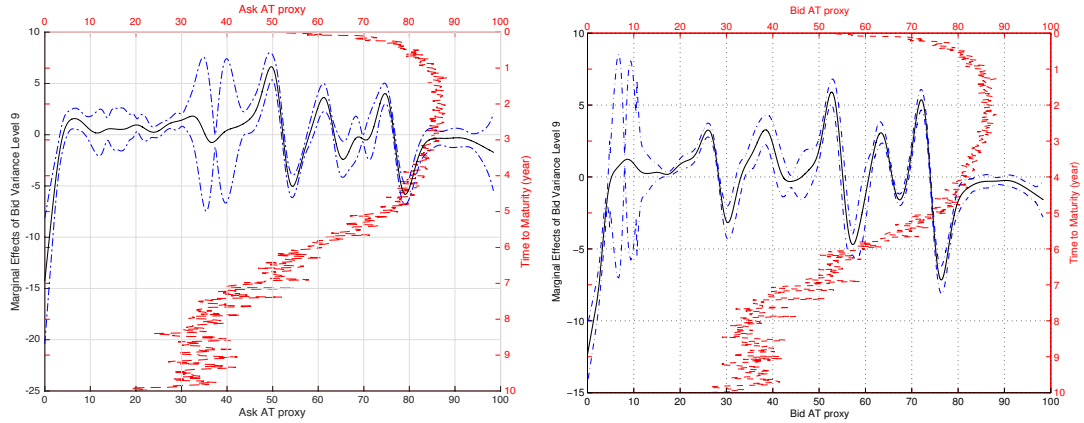
impacts coming from the bid side AT proportion, there is no significant effects on ask variance ratio when the fraction of bid side AT is below 51% and above 83%. However, the bid side AT proxy has positive effects on the ask variance ratio within a proportion of 51% to 56%, 59%to 63% and 67%to 74%. In these periods, algorithmic traders play a liquidity supplier role in the market, and result in the increasing both market liquidity and variance ratio.

Following each marginal effects increasing period, there is also a drop down period for the ask variance ratio in the rest proportion within 55% and 79%. The bid side AT proxy brings negative impacts on the ask variance ratio and results in dramatically decreases. The reason could be that the market reaches its saturation limits and orders from bid side ATs start to cancel each other out, therefore the liquidity contracts and the ask variance decreases. The effects from AT proportion on the bid variance ratio at level 9 display more fluctuations and less extreme than the ask variance ratio, and become stable above the proportion of 80% - 85% ATs. Subplots (e) and (f) describe that the fraction of AT has the ability to decouple between bids and asks of the book. The evidence indicates that ask side AT has limited impacts on the bid–ask covariance ratio at time scale level 9 before the proportion of AT below 42% and above 82% approximately. When the fraction of ask side AT is between 42% - 45%, 50% - 60% and 66% - 74%, the impact of AT on bid–ask covariance ratio is negative. This means the proportion of AT results in the bid–ask spreads decoupling. When the fraction of ask side AT is within 45% - 50% and 60% - 66%, the AT substantially increases the bid–ask covariance ratio, which indicates that the proportion of AT decreases the decoupling degrees of the bid–ask spreads. For the impacts from bid side AT activities, when the bid AT proxy increases to 40% - 52% and 62% - 68%, the marginal effect on bid–ask covariance ratio is substantially increased. Once the proportion of AT is between 52% - 62% and 68% - 74%, the bid–ask covariance



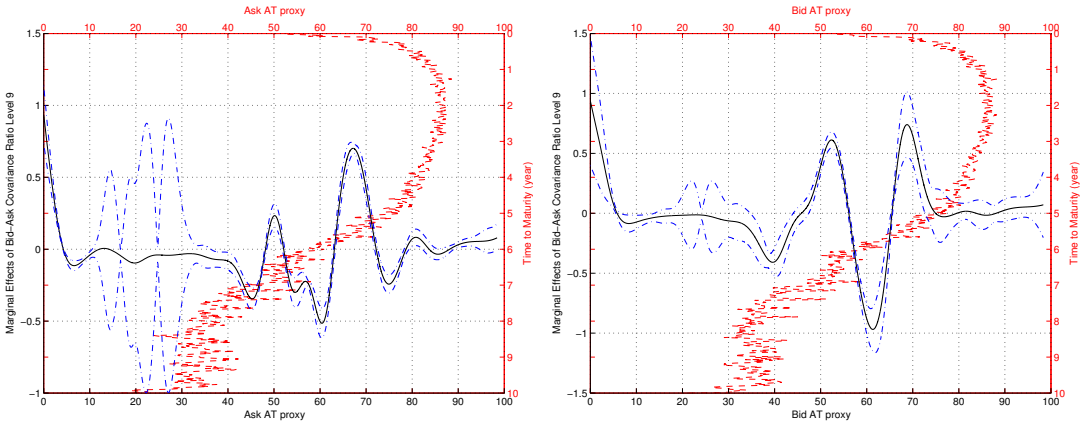
(a) Ask Variance Level 9 - Panel

(b) Ask Variance Level 9 - Panel



(c) Bid Variance Level 9 - Panel

(d) Bid Variance Level 9 - Panel



(e) Bid Ask Covariance Level 9 - Panel

(f) Bid Ask Covariance Level 9 - Panel

Figure 4.10: Marginal Effects of Variance and Covariance at Timescale Level 9 Thresholded at 200 ms

Notes: The red dashed line represents the evolution of daily AT proxy thresholded at 200 ms for order book level 2 in function of the average time to maturity. The black continuous line depicts the marginal effects of AT on variance ratio in Subplots (a) to (d) and, on covariance ratio in Subplots (e) and (f), both at timescale 9 with the panel dataset. The blue dashed lines are the 95% lower and upper confidence bounds of marginal effects.

ratio dips and results in the increasing degrees of bid–ask decoupling.

4.4 Chapter Summary

This chapter has provided a comprehensive empirical analysis of the relations between algorithmic trading at high frequencies with execution risks on the Eurodollar Futures markets. I have analyzed the entire limited order book and transactions of this market from 2008 to 2014, then estimated the marginal effects of the HFT activities on a series of market quality indicators using a novel semi-parametric partially linear approach. The market quality indicators contain both nine different quotes variance-covariance ratios and eight different market liquidity spreads, which have been measured in Chapter 2 and Chapter 3. I also adopted the AT proxies (the fraction of algorithmic trading thresholded under 25 to 200 milliseconds) as the dependent variable from Chapter 3.

I find that the AT proxy does impact various market quality measures, but overall it is not in a direction that will fully satisfy the detractors of high speed automated quoting or its backers. Therefore, I demonstrate that there are critical levels of the saturation of algorithmic trading in terms of the impact of HFTs on market quality. This may explain the contradictory evidence found in prior studies of this type. Specifically, a lower proportion of HF algorithmic trading has a detrimental marginal effect on market quality, a relative higher fractions of AT has a positive marginal effect, and once the AT proxy passes a certain saturation point the marginal effect disappears. I also have applied a large number of robustness checks by alternating the specification and applying differing instrumental variables in the semi-parametric analysis. In addition, I decompose the analysis by delivery date and found that my results are not a statistical artefact.

I have illustrated that the direction and non-linear nature to the effects are predicted by a standard adjustment to a model of asymmetric information and trading capacity within a standard rational expectations framework. Hence, whilst the results maybe surprising from the viewpoint of a linear analysis, they have a sound theoretical explanation based on expected trading patterns. There is substantial evidence that at long maturities high frequency transacting behavior is associated with deteriorations in market quality. However, when the majority of traders are operating in this manner, their impact is at worst insignificant, but can improve liquidity, by reducing a variety of measures of market illiquidity and increasing market depth. One significant result is, however, that for measures of correlation scaling, high frequency trading, at critical saturation points, is associated with a significant degree of decoupling between inside best bids and asks, but this is not a ubiquitous finding and indeed at points the opposite relationship is observed.

Chapter 5

High Frequency Price Discovery and Price Efficiency on Interest Rate Futures

In this chapter I examine the impact of price efficiency of high-frequency trading in the Eurodollar futures market. In a high-frequency world, the price can be decomposed into the efficient price (long term) and the pricing error (short term). To capture price efficiency, I calculate the mid-price return autocorrelation following the intervals: tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec. Then, I utilize a vector autoregression to estimate the pricing error as the deviation of transaction prices from the efficient prices to illustrate the structure of these costs. I have constructed a unique dataset using the complete messaging history (quotes and transactions) for the Eurodollar futures limit order book from 2008 to 2014. The findings suggest that the mid-quoted return autocorrelations are positive and gradually increase from the shortest time interval to the longest time interval. The adjustment time of trade returns reverts to equilibrium in a very short time interval of one second or less. Considering the maturity effect in the futures market, I find that the trade prices are less sensitive to incorporating any available information into the market as the Eurodollar future approaches its maturity.

5.1 Introduction

The speed of access to financial markets through electronic communication networks (ECNs) and the frequency of updating from automated algorithms is increasing dramatically. For instance, in Eurodollar future market, the average updating speeds has dropped from about 20 minutes to 600 milliseconds (ms) from 1996 to 2014. Futures contracts have always been active and liquid markets; however, their microstructure and the subsequent understanding of how prices are formed have only recently been brought into focus. This chapter provides a comprehensive empirical microstructure analysis to answer how fast you need to be in the Eurodollar future market. My approach is to apply a set of empirical measurements to evaluate both quotes and trades efficient reaction time in this Eurodollar future market. I demonstrate that the trading reaction time is below one second, for the contracts close to maturity, the trading reaction time even within 100 milliseconds. Furthermore, the quoting reaction time is faster than trades. For instance, the mid-quote can incorporate new information and adjust itself to a relative efficient price with 15 milliseconds.

This chapter looks at a very large market for interest rate derivatives, the Eurodollar futures market on Chicago Mercantile Exchange, and undertakes an extensive analysis at a very high-frequency domain. The empirical contribution of this chapter is threefold. First, I document the high-frequency price formation mechanism for the ED market and determine the effectiveness of various trade direction indicators (including a new version designed specifically for this market) in determining the future direction of prices. Second, I utilize information from the complete limit order book to assess the impact of high-frequency transactions on the instantaneous level of liquidity for the market. Finally, I document the degree of predictability inherent in the order-book as shocks to order-flow permeate

across the term structure and through time.

Comprehensive analytical work on data sets of over a billion observations are still rare in financial economics and econometrics. This chapter presents one of the first studies of this type on the most actively traded of markets, futures, in particular futures on interest rates. When working in this domain, even simple tasks, such as computing auto-correlations requires careful implementation and, in part, the major innovation of this work is in carefully documenting the appropriate strategies for this type of analysis. Methodologically, I will also introduce a new trade direction indicator for use in adverse selection and intra-day price impact models. In many respects this is the key contribution as this trade direction indicator is fundamental to determining buy and sell trades with very high frequency data. To this end I will use a Volume Weighted Average Price (VWAP) from the entire limit order book (buy and sell side) and then identify each trade as a buy or sell side based on its volume weighted position in the order book.

With the progress of technology in the financial markets, the microstructure plays a starring role because of its faster trading speed. High-frequency trading causes some fundamental changes in asset pricing, especially in liquidity and the price discovery process (see O'Hara, 2003; Riordan and Storkenmaier, 2012). For instance, informed traders no longer only indicate those with access to private inside information and those who take the opportunity of arbitrage; in a high-frequency world, if some HFTs receive, react, and process information faster than other traders, they are also treated as informed traders. Hence, informed traders do not necessarily hold fundamental information, which can generate adverse selection and increase trading costs in the market for hedgers. Besides the informed trading, HFTs tend to submit large amounts of orders and then cancel them instead of executing them, which creates a phantom liquidity and

gives other traders a false impression about the market liquidity.

Since high-frequency traders can play the role of informed traders, this can result in asymmetric information, decrease the capacity of price to incorporate the available information, and eventually harm the market efficiency and increase trading costs, see Chapter 3 and 4 for discussion on this. However, some HFTs also act as liquidity providers, rather than relying on informed trading to gain profits. In this case, HFTs do have benefits for the market liquidity and decrease the trading costs. This is one of my motivations - to figure out the high-frequency price discovery in the Eurodollar future market. Furthermore, as one of the most liquid and actively traded contracts in global financial markets, the three-month Eurodollar futures contract does not get enough academic attention. I construct a unique limit order book dataset of Eurodollar futures from 2008 to 2014, involving $7.66 quintillion ask total volumes, $7.56 quintillion bid total volumes, and $2 quadrillion trade volumes.

The increasing margin costs and strict regulations on interest rate swaps will cause the significant transfer of trading activities from the opaque OTC-IRS to Eurodollar strips to construct synthetic swaps. The significance of the Eurodollar futures market is one reason I study the microstructure of this particular market. Additionally, in my Eurodollar futures dataset, 65% of total activity is from traders at high-frequency when measured by fraction of messages on the order book at 25ms or faster, a fraction that is consistent with the evidence from CME Group [2010], where 64.46% message updates are from accounts registered as being automated trading accounts in the 2010 Eurodollar futures market. The large proportion of HFTs among the total traders is another reason this study focuses on the high-frequency price discovery process in the Eurodollar futures market.

Pricing models of the type used in this chapter decompose the observed price into the efficient price (long term) and the pricing error (short term fluctuation from the efficient price). As previously noted, the absolute degree of autocorrelation in the mid-price returns provides evidence on the level of price efficiency. Higher absolute magnitudes in returns indicate persistent fluctuations away from the mid price. In this chapter I provide evidence of the price efficiency from the mid-price return autocorrelations following a full range of time intervals: tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec. I illustrate this approach using the complete messaging history (quoted in five levels and transactions) for the Eurodollar futures limit order book from 2008 to 2014. The findings indicate that the mid-quoted return autocorrelations are positive and gradually increase from the shortest time interval to the longest time interval, except for the negative return autocorrelation at the tick level. Compared with other studies in the same field (see Anderson et al., 2013; Chakrabarty et al., 2014; Hendershott and Jones, 2005), most of them only measure quoted return autocorrelations in seconds or minutes, even at a daily level, because of the challenge to process massive quantities of microstructure data and the lack of clean processed data from the equity market.

Several interesting stylized facts emerge from preliminary analysis of this type of order-book data. One important implication of the decomposition of autocorrelation is for standard trade direction algorithms such as the Lee-Ready trade direction indicator. Standard implementations appear to give erroneous results (for instance all the trades are classified as buys or all sells for an entire day). Hence, one of the major contributions of this chapter is to introduce a new trade-direction indicator that takes advantage of the complete limit order book and

its statistical properties. I have developed a new volume weighted average trade classification algorithm specially designed for the Eurodollar future market. The algorithm eliminates the iceberg trades and utilizes the volume weighted average price over five levels of Eurodollar limit order book to classify trades directions. The efficiency rate of this algorithm, 95.88%, is three times efficient as the traditional Lee-Ready algorithm using the Eurodollar futures (see Table 5.1).

To measure the deviation of the trade prices from the actual efficient prices, I implement the vector autoregression (VAR) to estimate the pricing error and examine the impulse responses of trade return to trade directions, trade volume, square root of trade volume, and ask and bid market concentrations[1]. The responses imply that all factors have an impact on trade returns and the speed of the trade return adjustment is very fast, occurring within one second. This also indicates that the price discovery studies need to be conducted within very short time intervals due to the current high-speed trading. Finally, one of the major differences between equity assets and future assets is that the future has its own time to mature. Therefore, I conduct the research to verify the effects of the days to maturity on the price discovery and market quality indicators for the Eurodollar futures contracts.

The remainder of this chapter is structured as follows: Section 5.2 reviews the relevant high-speed market microstructure literature, especially regarding the high-frequency price discovery. Section 5.3 presents varying methodologies and the construction of variables to examine price discovery in the Eurodollar market. For instance, I use the price and return autocorrelations as measurements for price efficiency, and then decompose transaction prices and employ a vector au-

[1] As the number of buyers and sellers is significant for shifting trade prices, I consider the signed bid-side and ask-side market concentrations as dependents on price discovery mechanisms via the VAR model and impulse response functions

toregression model to proxy the pricing errors as the deviation of the transaction price from the efficient price. Following the VAR model, I also examine the trade price dynamics via impulse response functions and then utilize a multivariate linear regression to estimate the impact from the days of maturity of the Eurodollar market microstructure. Section 5.4 describes the data with descriptive statistics and focuses on the results interpretation with the price discovery and market quality of the Eurodollar future markets. Section 5.5 summarizes the implications of my findings in general and provides some comments.

5.2 The Dynamics of Price Discovery

Market liquidity and market informativeness are the fundamental attributes of financial markets. The purpose of buying and selling in a continuous limit order book, such the Globex platform, is to provide a mechanism for processing information and subsequently to efficiently value the assets in question. As one of the desirable features of financial markets, market liquidity is the ability to purchase or sell a substantial amount of financial assets without causing the dramatic shifting of asset prices. Bessembinder and Venkataraman [2010] and Hendershott et al. [2011] provide the measurements of a range of common liquidity spreads, including bid-ask spreads, quoted depth, effective spreads, realized spreads, and adverse selection spreads. Whether the increasing proportion of HFTs does contribute to the market liquidity or damage the liquidity in the financial markets, it has recently been brought into focus, especially after the 2010 Flash Crash in the E-mini S&P 500 future market. Jovanovic and Menkveld [2010], Hendershott et al. [2011], Carrion [2013], and Menkveld and Zoican [2014] demonstrate that HFT can narrow bid-ask spreads, increase liquidity, reduce trading costs, and

make the market more efficient. Whereas, HFT can access information faster than other traders, which could result in the asymmetric information across market participants. Brogaard et al. [2014] propose that the HFTs generate adverse selection costs to other investors, associated with the reduction of market liquidity and the increasing transaction costs, because the HFTs hold better information than other investors. In addition to the liquidity effects on asset prices, the rest of this section will introduce and focus on another function of markets on asset price behavior, the dynamics of price discovery, and the effects of HFTs on the price efficiency in the financial markets.

The price discovery function is the process of determining a price that balances the supply and demand of futures contracts. Prior research uses the term 'price discovery' to describe a variety of empirical characteristics. For instance, price discovery can be considered as a process of finding an equilibrium price [Schreiber and Schwartz, 1986]. Hasbrouck [1995] provides the definition the ability of financial markets to impound new information. Similarly, Baillie et al. [2002] also mention that price discovery reveals the relation between prices and information from one or multiple markets. Lehmann [2002] describes price discovery as the ability to efficiently incorporating information (learning from investors trading behavior) into the market prices. These explanations suggest that price discovery is a dynamic process to reach a state of equilibrium with the rapidly adjusting market prices to replace the old equilibrium with the new one through new information.

Prior studies mainly utilize four types of measurements to interpret the efficiency of price discovery: the return autocorrelations (see Anderson et al., 2013; Eom et al., 2004; Foley and Putniņš, 2014; Hendershott and Jones, 2005), the variance ratio (see Castura et al., 2010; Lo and MacKinlay, 1988), the price delay

(see Hou and Moskowitz, 2005), and pricing error (see Hasbrouck, 1993). This chapter employs both the autocorrelation and the price error methods to assess the efficiency of price discovery in the Eurodollar future market and is the first to do so on ultra-high frequency data.

The first method, the autocorrelation of mid-price returns within a certain time interval, can reflect the degree of quotes prices deviating from the random walk, and imply the predictability of short-term returns (see Anderson et al., 2013; Foley and Putniņš, 2014; Hendershott and Jones, 2005). Small return autocorrelation indicates the better price discovery capacity of this market. Hendershott and Jones [2005] employ the mid-quote return autocorrelation approach to assess the association between price efficiency and market transparency, using both transactions and quotes data of the three U.S. exchange-traded funds (ETFs) from August 2002 to October 2002. They state that when the Island electronic communications network stops displaying the limit order book and goes dark, the ETF market reduces the transparency, decreases the efficiency of price discovery and worsens trading costs. Once the Island redisplays its limit order book, price efficiency increases, and trading costs drop. Anderson et al. [2013] demonstrate that the return autocorrelation has the explanation power for the efficiency of price adjustment. They break down the return autocorrelation into four components: the partial price adjustment, the non-synchronous trading effect, the bid-ask bounce and the time-varying risk premia. Using 16 years of TAQ data from 1993 to 2008 and eight groups within a two-year sub-period, they calculate the daily return autocorrelations and suggest that the partial price adjustment to new information is the main contributor to the predictability of a return autocorrelation.

Consistent with the Hendershott and Jones [2005] study, Foley and Putniņš [2014]

also measure the absolute mid-quote return autocorrelation to investigate the effects of dark trading on the Canadian markets quality, where a minimum price improvement rule has been approved. In addition to the absolute return autocorrelation approach, Foley and Putniņš [2014] conduct an experiment adopting the variance ratios, high-frequency volatility and delays in reflecting market-wide information as proxies of information efficiency. Using the tick-by-tick Toronto Stock Exchange (TSX) Composite Index data in Canada between August 2012 and December 2012, they suggest that the dark trading on limited orders has a positive impact on information efficiency, which indicates prices are closer to follow the random walk under the assumption of the information efficiency theory.

The second approach utilized in this chapter is the Hasbrouck's vector autoregression model to estimate the empirical pricing error. Hasbrouck [1991a,b, 1993] introduce a vector autoregression and vector moving average (VMA) framework to measure the price errors as an indicator of price discovery. Hasbrouck [1991a] proposes a new measurement of trades information based on the variance decomposition of efficient prices, where prices can be separated into the trade-correlated components and trade-uncorrelated components. Hasbrouck [1991a] also points out that trade direction indicator is also associated with the changes in quoted prices. Furthermore, Hasbrouck [1993] decomposes the trade prices into a random walk component and a stationary stochastic component.

The first random walk component can be the indicator for an efficient price following the information efficiency assumption, and the second stochastic component measures the deviations from the efficient prices, which is also called as a pricing error. He points out the standard deviation of pricing error as a proxy to indicate the distance between the effective price and the actual trade price. Hence, the lower standard deviation of the pricing error, the more efficiency of price discovery

process in the market. If the pricing error is high, then they observed transaction price would be considered to be far away from the efficient price. Using a VARMA model with a sample of NYSE stocks order flow information, Hasbrouck [1993] considers the relation between trade returns and the trade classification indicators, and compute the average pricing error standard deviation is around 0.33 in the NYSE equity market.

The VAR model of Hasbrouck [1991a,b, 1993] has been widely utilized to assess the price discovery in microstructure studies (see Barclay and Hendershott, 2003; Boehmer and Kelley, 2009; Boehmer and Wu, 2013; Chakrabarty et al., 2014; Dufour and Engle, 2000; Engle and Patton, 2004; Krishnamurti et al., 2003). Dufour and Engle [2000] utilize this VAR technique in the equity markets, and demonstrate that the increasing trading speed enhances the ability of price adjustment to new information. To investigate the price discovery of trades after hours, Barclay and Hendershott [2003] employ the vector autoregression techniques to decompose the information into public and private components in the Nasdaq equity market. Krishnamurti et al. [2003] adopt the Hasbrouck's VAR model and the pricing error approach to compare the market quality of two major Indian stock exchanges with the different governance structures. Their finding suggests that demutualized stock exchanges have a better market quality with the lower average standard deviation of pricing error and the more efficiency price adjustment. Boehmer and Kelley [2009] and Boehmer and Wu [2013] also provide empirical evidence with the Hasbrouck pricing error method to assess the quality of the NYSE equity market in the market microstructure area.

In the high-frequency market microstructure domain, the market is fundamentally different from the low frequency world. A serial empirical studies are focusing on the high-frequency algorithm trading and the price efficiency in financial markets

(see Brogaard et al., 2014; Castura et al., 2010; Chaboud et al., 2014; Frino and McKenzie, 2002; Grünbichler et al., 1994; Hendershott and Riordan, 2011; Huang and Stoll, 1994). Grünbichler et al. [1994], Huang and Stoll [1994] and Frino and McKenzie [2002] devote to the impacts of the screen trading and suggest it benefits the efficiency of the prices. Grünbichler et al. [1994] explore the relation between the price formation and the speed of trading utilizing the German stock index (DAX). Because the underlying asset and the future asset of DAX index are traded in different trading platforms, they find the existence of a lead-lag relation between the spot and future prices for DAX index. The results suggest that the reaction of the screen traded future prices to information are 15-20 minutes faster than the spot prices of floor traded DAX index. Therefore, the screen trading increases the price efficiency and accelerate the price discovery process. The lead-lag relation between the spot and future market has been supported by Harris [1989], Stoll and Whaley [1990], Chan [1992], Huang and Stoll [1994], and Frino and McKenzie [2002]. Consistent with Grünbichler et al. [1994] study, Frino and McKenzie [2002], employing both screen-traded cash and future markets of FTSE 100 index, put forward that the screen trading can enhance the price discovery in both spot and future markets.

The impact of high-speed algorithmic trading on the incorporation of information into asset prices has been widely studied (see Brogaard et al., 2014; Castura et al., 2010; Chaboud et al., 2014; Hendershott and Riordan, 2011. Castura et al. [2010] adopt the variance ratio method from Lo and MacKinlay [1988] to test the effects of HFT on price efficiency. Variance ratio refers to the price variance at different time intervals. The theoretical expectation of a variance ratio means that the variance at various intervals should follow a linear relation in an efficient market. Castura et al. [2010] compute 1-, 10-, 60- and 600-second variance ratios using the Russell 1000 and Russell 2000 between 2006 and 2009. The findings suggest that

the increase of HFT activities has benefited the U.S. equity markets efficiency.

Utilizing DAX 30 stocks as a case study on the Deutsche Boerse, Hendershott and Riordan [2011] conclude that the algorithmic traders can efficiently place quotes and have an influence in adjusting the liquidity to bring the price closer to the efficient price. This indicates that the algorithmic traders have the impacts on improving market liquidity and enhancing quote efficiency. Additionally, Chaboud et al. [2014] analyse the relation between algorithmic trading and price discovery process with a comprehensive dataset in the foreign exchange market, Euro-Dollar, Dollar-Yen, and Euro-Yen from 2003 to 2007. Using a structure vector autoregression model, they suggest that the algorithmic trading is associated with the reducing frequency of arbitrage opportunities and the reduction in the autocorrelation of high-frequency returns. Hence, the high-speed of algorithmic trading improves the price discovery process and the information efficiency in the market.

Recent work by Chakrabarty et al. [2014] and Brogaard et al. [2014] point out some mixed empirical results on the impacts on the price formation and market qualities from the HFT and its trading strategies. Brogaard et al. [2014] suggest that HFT trading strategies have a multitude of effects on price efficiency.

If prices are separated into the efficient prices and the pricing errors, and the pricing errors can be viewed as an implicit trading cost to long-term investors then the HFT trade direction is consistent with the pricing error. HFT will subsequently play a role in increasing the trading cost for long-term investors. In contrast, if HFT traders go against this pricing error, HFT can be treated as decreasing trading costs for investors and providing liquidity to the market. By controlling the trading speeds, Chakrabarty et al. [2014] compare the equity market quality between the Securities and Exchange Commission (SEC) pre-ban

and post-ban period. Their findings suggest that lowering down the trading speed has no effects on the long-term price efficiency, but the price efficiency at short timeframe decreases at the post-ban period. In other words, the price formation within short period can be benefit from the high-frequency trading. However, Chakrabarty et al. [2014] find little evidence of a deleterious effects of banning naked access to US equity exchanges by the SEC and cite serval positive outcomes from this regulatory action, including reducing the adverse selection effects and trading costs.

Following the Hendershott and Jones [2005] autocorrelation and the VAR model of Hasbrouck [1991a,b, 1993], I assess the information incorporated in the order flow and trade prices. This chapter not just includes the trades return autocorrelation, but also a battery of the quote price autocorrelations from tick-by-tick interval to half hour interval. For the VAR model and pricing error approach in this chapter, I also measure the effect of a shock or impulse response to the trading process. Moreover, this chapter includes the trade direction indicator and trade volume, and also adds bid-side market concentration and ask-side market concentration, as factors in the VAR model. The next section will describe these measurements in detail.

5.3 Empirical Methodology

My analysis is in four parts. First, using high-frequency limit order book data, I generate a series of mid-price autocorrelation and mid-price return autocorrelation at different time intervals. Second, I introduce a new volume weighted average trade classification algorithm, and then utilize it to calculate multiple market liquidity spreads and a series of spreads autocorrelation at varying time

intervals. Third, I use the vector autoregression and the impulse response function to proxy the tick-by-tick pricing error to assess the trade price dynamics. Finally, I use multivariate linear regression to estimate the impacts of the distance to maturity on price discovery and price efficiency. To maintain consistency, there are two notational conventions for the time. One is the daily time index, denoted as subscript t, where $t \in \{t, \ldots, T\}$, $t \in \mathbb{Z}_+$. The other is the subscript $k \in \{1, \ldots, K\}$ is the intraday time index to deal with the high-frequency intraday transactions and limit order book data, where $k \in \mathbb{Z}_+$. Individually, each of these steps is not methodologically new, however, their use with a dataset that has well over a billion observations, presents many challenges in terms of implementation. In the online appendix, I document the various strategies needed to overcome these challenges. This, in part, is the core novelty of the research.

5.3.1 Price Efficiency

One of the most common approaches in market microstructure is documenting return autocorrelations. Classical representations of market efficiency suggest that when information is efficiently incorporated into asset prices then autocorrelations in returns will disappear rapidly. However, when the trading mechanism prevents the efficient incorporation of new information into observed prices, it is likely that these prices will deviate from the random walk. Hence, the return autocorrelation can be an indicator to measure the degree of quotes prices against the random walk. Eom et al. [2004], Hendershott and Jones [2005], Anderson et al. [2013], and Foley and Putniņš [2014] employ the mid-price return autocorrelation at varying time interval to determine the predictability of short-term return.

In this chapter I will estimate models directly onto the trade and quote data for the market at high frequency, in contrast to the daily aggregate measures used in the preceeding chapters. The most challenging intra-day regressions I have undertaken up to now (for instance the wavelet analysis) were estimated for a single days activity. For computational reasons I have decided to utilize a linear parametric representation for the price process and then analyze the direction of returns at various frequencies. It is now important to define what is meant by the 'price process'. In this case I will study the mid-price return, denoted $r_{\tau k}$. Adjustments to this price are formed by adjustments in the price and volume within the quotes forming the limit-order book of the market. Unlike trade ticks, this price is update at an almost continuous level and is as close to a continuous diffusion as is possible in this context.

Let the intraday time interval be denoted by τ, for $\tau \in \{$tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec$\}$. Additionally, I measure both the volume weighted average mid-price returns at diverse time intervals τ, denoted as $r^{v}_{\tau k}$, and mid-price return for each order book level $j \in \{1, 2, \ldots, J\}$ at diverse time intervals τ, denoted as $r_{j\tau k}$. The mid-quote return autocorrelations can be expressed as:

$$\Omega^{R}_{\tau\eta k} = \rho(r_{\tau k}, r_{\tau k-\eta}) \tag{5.1}$$

where Ω^{R} stands for the autocorrelation of mid-price returns, ρ is the autocorrelation function, r is the logarithm of price changes, $r_{\tau k} \in \{r^{v}_{\tau k}, r_{j\tau k}\}$, τ is the length of mid-price returns time intervals, k is the intraday time index, η is the orders of autocorrelations $\eta \in \{1^{st}, 5^{th}, 10^{th}\}$, and j is the level of the order book. I estimate the volume-weighted average autocorrelation across the five levels order

book and the autocorrelation of each level order book (level one to level five). To estimate the daily mid-quoted return autocorrelation, I calculate the first-order, fifth-order, and tenth-order autocorrelations, and report the daily mean value and median value.

Unlike other studies, I do not take the absolute value of the autocorrelation yields, because I will use the sign of the autocorrelation as an extra piece of information to be incorporated into prices. Indeed, because I observe many negative autocorrelations are the very highest frequency, this inversion is deemed to be important. As such, I calculate the bid price, ask price, and mid price autocorrelations at the tick-by-tick time interval

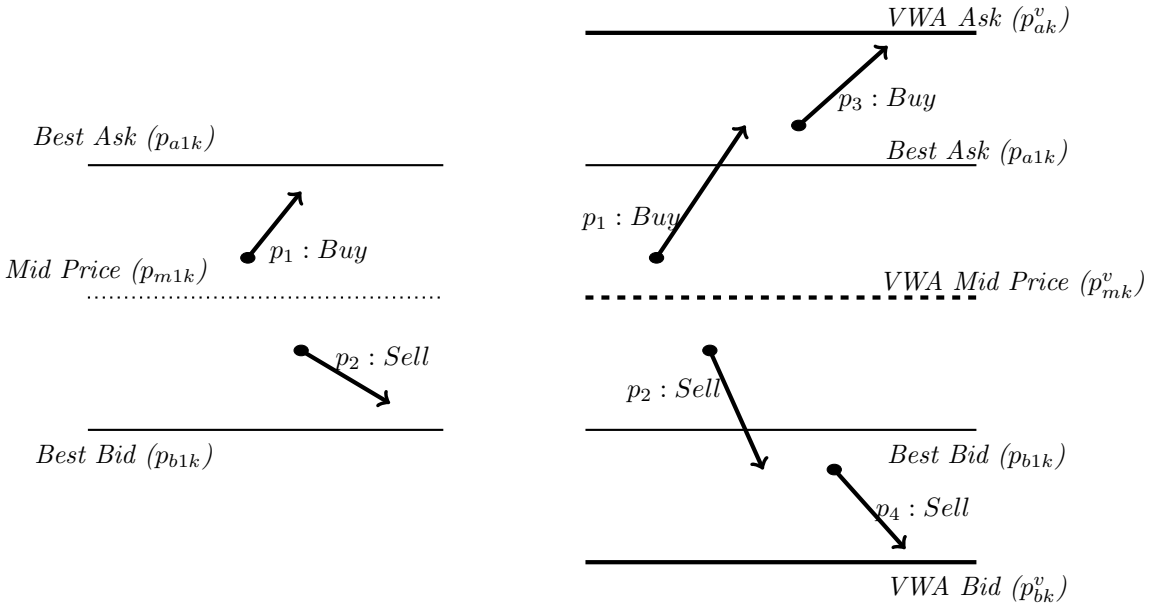
$$\Omega^{P}_{\eta k} = \rho(p_k, p_{k-\eta}) \tag{5.2}$$

where Ω^P is the autocorrelation of prices, ρ is the autocorrelation function, η is the orders of autocorrelations $\eta \in \{1^{st}, 5^{th}, 10^{th}\}$, k is the intraday time index, the price $p \in \{p^v_a, p_{aj}, p^v_b, p_{bj}, p^v_m, p_{mj}\}$ involves ask, bid, and mid prices, p^v_a is the volume weighted average ask prices over the five levels, p_{aj} is the ask prices for each level $j \in \{1, 2, \ldots, J\}$ in the limit order book, p^v_b is the volume weighted average bid prices, p_{bj} is the bid prices for each level j, p^v_m is the volume weighted average mid prices, and p_{mj} is the mid prices for each level j. This chapter explores the first-order, fifth-order, and tenth-order bid, ask, and mid-quote autocorrelations using both the volume weighted average prices and quotes in the level one to level five market data in the limit order book, and report the daily mean value and median value.

5.3.2 Market Quality Indicators

Following the comprehensive liquidity measures provided by Hendershott et al. [2011], I generate different market liquidity spreads ($\tilde{\mathscr{S}}$) and spreads autocorrelation ($\Omega^{\tilde{\mathscr{S}}}$) then use these indicators as dependent variables in the regression, including bid-ask spreads ($\tilde{\mathscr{S}}^{Q}$), quoted half-spread ($\tilde{\mathscr{S}}^{Q^{1/2}}$), quoted depth ($\tilde{\mathscr{S}}^{D}$), effective half-spreads ($\tilde{\mathscr{S}}^{E}$), realized spreads ($\tilde{\mathscr{S}}^{R}$), and adverse selection ($\tilde{\mathscr{S}}^{AS}$). Both realized spreads and adverse selection are measured at varying 26 time intervals from tick-by-tick to 30 minutes. It is worth to mention that I alter the Lee-Ready algorithm [Lee and Ready, 1991] using the volume-weighted average price across five levels order book, because of the inaccuracy of the Lee-Ready algorithm for high-frequency data.

O'Hara [2015] outlines that the modalities of market making and trading activities is changing rapidly and quite radically. The low-latency trading could fundamentally change the rules in the market microstructure world. Some existing basic microstructure models and approaches are not able to efficiently capture the market dynamics via the high-speed trading, such as the Lee-Ready algorithm trader classification. Figure 5.1 compares the volume weighted average trade classification algorithm with the Lee-Ready algorithm. Previous studies are likely to employ the inside quotes data, hence in Figure 5.1 Subplot (a) I use the first level limit order book data (best-bid best-ask in the equity markets) to illustrate the Lee-Ready algorithm, namely as the p_{a1k} and p_{b1k}. Lee and Ready [1991] propose that if a trade price (p) is higher (lower) than the midpoint of the inside spreads p_{m1k}, and then this trade belongs to a buyer-side (seller-side) initiated trade. If a trade price equals to the mid-quoted price, then the trade classification depends on the tick rule and the comparison with the previous trade price. In this case, if the trade price is above (below) its previous trade price, then it is a buy-side



(a) Lee-Ready Algorithm (b) Volume Weighted Average Trade Classification

Figure 5.1: Trade Classification Algorithms

Notes: This figure represents the traditional Lee-Ready trade classification algorithm (Subplot (a)), and the modified volume weighted average trade classification method (Subplot (b)). Assuming bid and ask prices are constant and follows the independent Poisson processes. Trade prices p_1, p_2 arrival within the inside quotes and t prices p_3, p_4 are outside the inside quotes. In Subplot (a), I only using the inside spread with its midpoints to illustrate the Lee-Ready algorithm, where the best ask and best bid are denoted as p_{a1k} and p_{b1k} respectively, and k is the intraday time index. The dotted line refers to the mid-price p_{m1k} of the inside spread. Subplot (b) depicts the volume weighted average trade classification method. First of all, the volume weighted average asks and bid prices across five levels limit order book can be expressed as: $p^v_{ak} = \frac{\sum_j p_{ajk} v_{ajk}}{\sum_{j=1}^{J} v_{ajk}}$ and $p^v_{bk} = \frac{\sum_j p_{bjk} v_{bjk}}{\sum_{j=1}^{J} v_{bjk}}$, where $j \in \{1, 2, \ldots, J\}$ is the j^{th} level of the order book; p_{ajk} and p_{bjk} are the ask price and bid price for level j at intraday time k; and v_{ajk} and v_{bjk} are the ask volume and bid volume for level j. The dashed line represents the volume weighted average mid-price p^v_{mk} which equals to the average of both p^v_{ak} and p^v_{bk}. There is a small fraction of trades locating outside the volume weighted average spreads. I eliminate these iceberg orders before classifying trades directions.

(sell-side) trade. Therefore, in Figure 5.1 Subplot (a), trade p_1 is a buyer-initiated transaction and trade p_2 is a seller-initiated transaction.

However, Lee and Ready [1991] suggest that their algorithm should implement with quotes and trade updating frequency above five seconds to have the accurate trade classification. As it is important to look deep into the tick level to understand the market microstructure world, I construct the Eurodollar contracts dataset in millisecond timestamps. Unlike other studies, I adopt all five levels order book data in this study so that one trade price corresponds with five different level mid-point prices. Therefore, the Lee-Ready algorithm is unable to handle my dataset.

To replace the Lee-Ready algorithm, this chapter adopts a modified version of trade classification algorithm, using the volume weighted average quotes and eliminating iceberg orders (see Figure 5.1 Subplot (b)). The volume weighted average trade classification algorithm can be measured in the following steps.

Step 1: Clean the Eurodollar dataset and eliminate bid prices p_{bjk}, ask prices p_{ajk}, and trade prices p_k containing NaN (not a number), zero, and infinity for each order book level.

Step 2: Match the timestamps k between trades and quotes to obtain the equal length of quotes and trades.

Step 3: Calculate the volume weighted average bid prices p^v_{bk}, ask prices p^v_{ak}, and mid prices p^v_{mk},

$$p^v_{ak} = \frac{\sum_j p_{ajk} v_{ajk}}{\sum_{j=1}^{J} v_{ajk}}, \quad p^v_{bk} = \frac{\sum_j p_{bjk} v_{bjk}}{\sum_{j=1}^{J} v_{bjk}}, \quad p^v_{mk} = (p^v_{ak} + p^v_{bk})/2 \tag{5.3}$$

where $j \in \{1, 2, \ldots, J\}$ is the j^{th} level of the order book; p^v_{ak} and p^v_{bk} are the

volume-weighted average ask price and bid price at the time of the k_{th} trade; p^v_{mk} is the volume-weighted average mid price at the time of the k_{th} trade; p_{ajk} and p_{bjk} are the ask price and bid price for level j at intraday time k; and v_{ajk} and v_{bjk} are the ask volume and bid volume for level j at intraday time k.

Step 4: Eliminate the iceberg orders and outliers. If any trade price is smaller than volume weighted average bids, $p_k < p^v_{bk}$, or larger than volume weighted average asks $p_k > p^v_{ak}$, I define it as an iceberg order. It locates outside the volume weighted average bid-ask spreads; hence I delete this transaction p_k and corresponding quotes (p^v_{bk}, p^v_{ak}, and p^v_{mk}) at the time k from the dataset.

Step 5: Calculate the buy-sell trade direction indicator q_k with the following expression:

$$q_k = sign(p_k - p^v_{mk}) \tag{5.4}$$

where $sign(.)$ is the signum function with the value of {-1, 0, 1}. Therefore, if the q_k equals to +1, it means the trade price p_k at time k is above the volume weighted average mid-price p^v_{mk} and belongs to a buyer-initiated trade. If the q_k equals to -1, it indicates the trade price p_k at time k is below the volume weighted average mid-price p^v_{mk} and belongs to a seller-initiated trade. If the q_k equals to 0, it is defined as no trades.

Following these rules, it is possible to see p_1 and p_3 are defined as the buy-side trades and p_2 and p_4 are the sell-side trades in Figure 5.1 Subplot (b). Table 5.1 demonstrates the comparison between the original Lee-Ready algorithm and the new volume weighted average trade classification algorithm. As one relatively active traded day, there are 7.1 million observations for five levels bids and asks and 29,793 trades on June 05, 2009 of GEH0 Eurodollar contract, which is around eight months before its settlement. Hence, the Lee-Ready algorithm can classify

Table 5.1: Comparison of Trade Classification Algorithms

GEH0 (June 05, 2009)	**Lee-Ready Algorithm**	**Volume Weighted Average Algorithm**
Total Asks	7,082,965	
Total Bids	7,082,965	
Total Trades	29,793	
Buy-side Trades (+1)	4,949	12,741
Sell-side Trades (−1)	4,805	15,824
Excluded Ask Side		98
Excluded Bid Side		205
Trades Direction Indicator (q)	**9,754**	**28,565**
Algorithm Efficiency Ratio	**32.74%**	**95.88%**

Notes: This table compares the trade classification by the Lee-Ready algorithm and the volume weighted average algorithm. I use the full order flow and executed prices of GEH0 on June 05, 2009 as a toy example to examine the reliability of both trade classification algorithms. Because of eight months before GEH0 maturity, this contract is relative active on this day with 29,793 trades observations. Total asks, total bids, and total trades are the total number observation of asks, bids, and transactions within one day. Buy-side and sell-side trades report how many trades are classified as buyer-side (+1) initiated trades or seller-side (−1) initiated trades. The volume weighted average trade classification method eliminates the iceberg trades as the outside the spread; therefore the number of iceberg trade is reported as the excluded ask side and excluded bid side. Trade direction indicators refer to the sum of buy-side and sell-side trades, and algorithm efficiency ratio is the number of trade direction indicator to the total trades.

9,754 trades with 4,949 buy-side trades and 4,805 sell-side trades. The efficiency ratio of Lee Ready algorithm is only around 32.74%, if I compute the ratio as the total number of trade direction indicator to the total number of trades. However, there are 28,565 trades defined by my new volume weighted average trade classification algorithm, with 12,741 buy-side transactions and 15,824 sell-side transactions (see Table 5.1). Additionally, the efficiency ratio of my algorithm is around 95.88%, which is three times higher than the Lee-Ready algorithm.

Besides, I also calculate both the daily average (dotted line) and the weekly moving average (black line) trade direction indicators across all Eurodollar contracts from 2008 to 2014 (see Figure 5.2). For each contract each day, the average of trade direction indicator can be expressed as $q_{it} = \sum_k q_{ik}/n_{it}$, where i is contract type, k is intraday time index, t is daily time index, n_{it} is the number of trades of contract i on day t. Then, I take the average value of daily trade direction indicator over 40 contracts, as $q_t = \frac{\sum_i \frac{\sum_k q_{ik}}{n_{it}}}{n_t^c}$, where n_t^c is the total number of contracts on day t. In Figure 5.2, if the trade direction indicator is below zero, it illustrates the ask-side trades are more than bid-side trades on that day. If the trade direction indicator is above zero, it means the amount of buyer-initiated trades is larger than sell-initiated trades. Overall, it seems like that the buyer-initiated trades and seller-initiated trades maintain the balance, except for the tendency of heavy seller-initiated trades during the period between March 2009 – August 2009 and May 2013 – October 2013.

After classifying the transaction using the new volume weighted average algorithm, I can measure the market liquidity spreads for each level or for the average across the whole limited order book. Denoted $\tilde{\mathscr{S}}$ is the set of liquidity spreads, where $\tilde{\mathscr{S}} \in \{\tilde{\mathscr{S}}^Q, \tilde{\mathscr{S}}^{Q^{1/2}}, \tilde{\mathscr{S}}^D, \tilde{\mathscr{S}}^E, \tilde{\mathscr{S}}_\tau^R, \tilde{\mathscr{S}}_\tau^{AS}\}$. The bid-ask spreads ($\tilde{\mathscr{S}}^Q$), quoted half spreads ($\tilde{\mathscr{S}}^{Q^{1/2}}$), quoted depth ($\tilde{\mathscr{S}}^D$) and effective spreads ($\tilde{\mathscr{S}}^E$) are measured at the tick-by-tick level, and both realized spreads ($\tilde{\mathscr{S}}_\tau^R$) and adverse selection indicators ($\tilde{\mathscr{S}}_\tau^{AS}$) are measured at varying 26 time intervals.

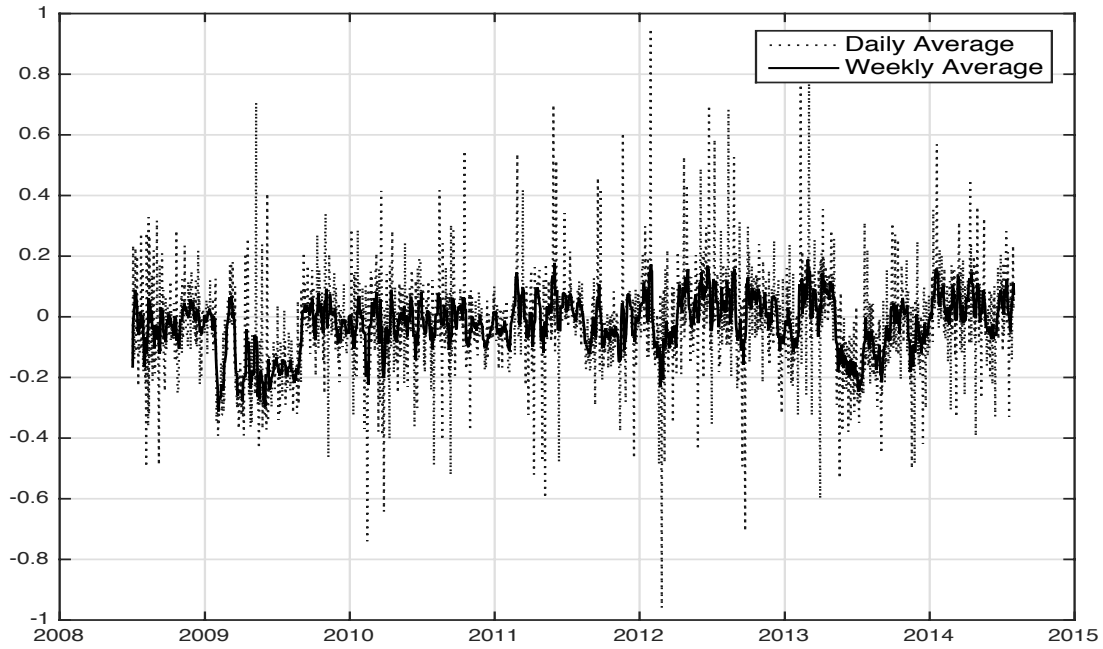


Figure 5.2: Average Trade Direction Indicator

Notes: This figure depicts both the daily average (dotted line) and the weekly moving average (black line) trade direction indicators across 40 contracts from 2008 to 2014. Using the volume weighted average trade classification algorithm, I classify each transaction as buyer-initiated trades (+1) and seller-initiated trades (-1) at tick level and denote as the trade direction indicator (q). The average of all 40 Eurodollar contract daily trade direction indicator is $\sum_i(\sum_k q_{ik}/n_{it})/n_t^c$, where i is contract type, k and t are intraday time index and daily time index, n_{it} is the number of trades of contract i on day t, and n_t^c is the number of contracts on day t.

Therefore, the market liquidity spreads can be expressed as follows

$$\tilde{\mathscr{S}}_{jk}^{Q} = 100(p_{ajk} - p_{bjk})/(p_{mjk}) \tag{5.5}$$

$$\tilde{\mathscr{S}}_{jk}^{Q^{1/2}} = 100(p_{ajk} - p_{bjk})/2(p_{mjk}) \tag{5.6}$$

$$\tilde{\mathscr{S}}_{jk}^{D} = p_{ajk}v_{ajk} + p_{bjk}v_{bjk} \tag{5.7}$$

$$\tilde{\mathscr{S}}_{jk}^{E} = q_{jk}(p_{jk} - p_{mjk})/p_{mjk} \tag{5.8}$$

$$\tilde{\mathscr{S}}_{j,k,\tau}^{R} = q_{jk}(p_{jk} - p_{mjk+\tau})/p_{mjk} \tag{5.9}$$

$$\tilde{\mathscr{S}}_{j,k,\tau}^{AS} = q_{jk}(p_{mjk+\tau} - p_{mjk})/p_{mjk}, \tag{5.10}$$

where $j \in \{1, 2, \ldots, J\}$ is the j^{th} level of the order book; k is an intraday time index; t is a daily time index; p_{ajk} and p_{bjk} are the ask price and bid price for level j at intraday time k, p_{mjk} is the quoted mid price for level j at the time of the k_{th} trade, p_{jk} is the actual transaction price for level j at the time of the k_{th} trade; v_{ajk} and v_{bjk} are the ask volume and bid volume for level j at intraday time k; q_{jk} is the buy-sell trade direction indicator, that +1 for buyer-initiated trades, -1 for seller-initiated trades, and 0 for no trades; and τ is the length of intraday time intervals, $\tau \in$ {tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec}.

I also consider liquidity spreads autocorrelations as indicators to assess the information efficiency with liquidity dynamics. The spreads autocorrelation can be computed as:

$$\Omega_{\eta k}^{\tilde{\mathscr{S}}} = \rho(\tilde{\mathscr{S}}_k, \tilde{\mathscr{S}}_{k-\eta}) \tag{5.11}$$

where $\Omega^{\tilde{\mathscr{S}}}$ is the liquidity spreads autocorrelations, η is the orders of autocorrelations $\eta \in \{1^{st}, 5^{th}, 10^{th}\}$, and $\tilde{\mathscr{S}}_k$ is the varying spreads at intraday time k and $\tilde{\mathscr{S}}_k \in \{\tilde{\mathscr{S}}_k^v, \tilde{\mathscr{S}}_{jk}\}$, $\tilde{\mathscr{S}}_k^v$ is the volume weighted average spreads autocorrelations at time k, and $\tilde{\mathscr{S}}_{jk}$ is the spreads autocorrelations in each order book level $j \in \{1, ..., 5\}$.

Besides these common liquidity indicators, I also consider other factors, including volatility, quote-to-trade ratio, bid-ask market concentration index, and the inverse of price. The daily price volatility, $\mathcal{V}$, is the daily high transaction price minus low transaction price, scaled by the high price, as defined in Hendershott et al. [2011]:

$$\mathcal{V}_t = p_t^h - p_t^l \tag{5.12}$$

where p_t^h is the daily highest trade price at day t, and p_t^l is the daily lowest trade price at day t. Moreover, the quote-to-trade ratio, $\mathcal{Q}$, also is an indicator of the ratio of the number of quotes to the total number of trades,

$$\mathcal{Q}_t = n_t^q / n_t \tag{5.13}$$

where n_t^q is the daily number of quotes at day t, and n is the daily number of trades.

The bid-side (ask-side) market concentration ratio, $\mathcal{C}_b$ ($\mathcal{C}_a$), is the volume of bids (asks) divided by the number of buyers (sellers), and can be estimated as

$$\mathcal{C}_{at} = \frac{v_{at}}{n_{at}}, \quad \mathcal{C}_{bt} = \frac{v_{bt}}{n_{bt}} \tag{5.14}$$

where v_{at} and v_{bt} are the daily total volume in ask-side and bid-side at day t, and the n_{at} and n_{bt} are the total daily number of sellers and buyers at day t.

5.3.3 Vector Autoregression Models of Pricing Errors

Under the random walk hypothesis, the observed transaction prices can be decomposed into random walk component and stationary component (see Hasbrouck, 1991a,b, 1993). The random walk component, m_k, is regarded as an indicator the efficient price based on the information efficiency theory, where the efficient price is the expectation of price incorporating all the information at time k. The stationary component, s_k, reflects the deviation of the trade prices from the efficient prices at time k, which is the pricing error. The higher pricing error is, the lower efficiency and less sensitivity of the transaction price absorbing new information.

Notationally, the observed logarithm of transaction prices is defined by:

$$p_k = m_k + s_k \tag{5.15}$$

The random-walk efficient prices, m_k, can be written as:

$$m_k = m_{k-1} + w_k \tag{5.16}$$

where w_k is the uncorrelated increments with $\mathbb{E}[w_k] = 0$, $\mathbb{E}[w_k^2] = \sigma_w^2$, $\mathbb{E}[w_k w_s] = 0$ for $k \neq s$, and $w_k \sim N(0, \sigma_w^2)$. The stationary pricing error component, s_k, is zero mean covariance-stationary process. Hasbrouck [1993] considers the microstructure effects of pricing error, and points out the pricing error includes an information correlated part, αw_k, and an information uncorrelated part, η_k, hence it decomposes into the following expression

$$s_k = \alpha w_k + \eta_k \tag{5.17}$$

where η_k is a disturbance uncorrelated with w_k. The information uncorrelated pricing errors is likely sourced from noise trading and inventory control impacts; and the information correlated component mainly comes from adverse selection effects and the lagging adjustment for new information. The return, r_k, is the difference between the logarithmic prices at time k and time $k-1$. The correspondence from Equation (5.15) to Equation (5.17) establishes that

$$r_k = p_k - p_{k-1} = (m_k + s_k) - (m_{k-1} + s_{k-1}) = w_k + s_k - s_{k-1} \tag{5.18}$$

where w_k and s_k are serially uncorrelated and r_k possesses nonzero auto-covariance at the first order only. The first-order moving average representation of return r_k

is

$$r_k = \epsilon_k - a\epsilon_{k-1} \tag{5.19}$$

Combined these two return expressions in Equation (5.18) and Equation (5.19), I now have the following relations:

$$\epsilon_k - a\epsilon_{k-1} = w_k + s_k - s_{k-1} \tag{5.20}$$

Under the interpretation of Beveridge and Nelson [1981] decomposition, the source for the pricing error, s_k is belongs to the information-correlated source, set $\eta_t = 0$ therefore

$$s_k = \alpha w_k \tag{5.21}$$

Therefore the equation (5.20) can also be represented as follow:

$$\epsilon_k - a\epsilon_{k-1} = (1+\alpha)w_k + \alpha w_{k-1} \tag{5.22}$$

where this equation also can split into two parts: $\epsilon_k = (1+\alpha)w_k$ and $a\epsilon_{k-1} = \alpha w_{k-1}$. Hence, the following identities can be verified

$$\alpha = \frac{a}{1-a} \tag{5.23}$$

$$w_k = (1-a)\epsilon_k \tag{5.24}$$

$$\sigma_w^2 = (1-a)^2\sigma_\epsilon^2 \tag{5.25}$$

$$\sigma_s^2 = a^2\sigma_\epsilon^2 \tag{5.26}$$

To explain the dynamic of trade prices and estimate the pricing error, I follow the Hasbrouck [1993] VAR framework on the tick-by-tick basis to compute the variance of the pricing error. Unlike the Hasbrouck's univariate autoregression, this chapter considers the dynamic relation among six factors: trade returns (r),

signed trade direction indicator (q), signed trade volume (v), signed square root of trade volume ($v^{1/2}$), signed ask-side market concentration ($\mathcal{C}_a$), and signed bid-side market concentration ($\mathcal{C}_b$), respectively.

The basic λ-lag vector autoregression model can be written as follow:

$$Y_k = c + B_1 Y_{k-1} + B_2 Y_{k-2} \cdots + B_\lambda Y_{k-\lambda} + u_k \tag{5.27}$$

where $Y_k = [r_k, q_k, v_k, v_k^{1/2}, \mathcal{C}_{ak}, \mathcal{C}_{bk}]'$ denotes as the vector-valued time series variables; k is the intraday time index; c is the constant matrix; B is the coefficient matrices and u is an error term. The concise matrix form of VAR model is

$$Y = \mathbf{B}X + U \tag{5.28}$$

where the dependent variables vector Y is $Y_k = [r_k, q_k, v_k, v_k^{1/2}, \mathcal{C}_{ak}, \mathcal{C}_{bk}]'$, the lagged set of explanatory variables X is $X = [Y_{k-1}, Y_{k-2}, \ldots, Y_{k-\lambda}]$, and $\mathbf{B}$ is the coefficient matrices of the lagged terms with

$$\mathbf{B} = [c, B_1, B_2, \ldots, B_\lambda] = \begin{bmatrix} c_1 & b_{1,1}^1 & b_{1,2}^1 & \ldots & b_{1,1}^\lambda & b_{1,2}^\lambda & \ldots & b_{1,6}^\lambda \\ c_2 & b_{2,1}^1 & b_{2,2}^1 & \ldots & b_{2,1}^\lambda & b_{2,2}^\lambda & \ldots & b_{2,6}^\lambda \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c_6 & b_{6,1}^1 & b_{6,2}^1 & \ldots & b_{6,1}^\lambda & b_{6,2}^\lambda & \ldots & b_{6,6}^\lambda \end{bmatrix} \tag{5.29}$$

and U is the white noise disturbances matrix with $u_k = [u_{1,k}, u_{2,k}, \ldots, u_{6,k}]'$ and $u_k \sim N(0, \sigma^2)$.

Any VAR(λ) can be rewritten as a VAR(1) with companion matrix representing the vector projection. Following the description of Hamilton [1994], let the

centered VAR system of equations be (5.27) be given by:

$$Y_k - \mu = c + B_1(Y_{k-1} - \mu) + B_2(Y_{k-2} - \mu) \cdots + B_\lambda(Y_{k-\lambda} - \mu) + u_k \tag{5.30}$$

hence, the VAR(1) format can be compactly rewritten in companion form as follows:

$$\hat{Y}_k = F\hat{Y}_{k-1} + V_k \tag{5.31}$$

where $\hat{Y}_k = [Y_k - \mu, Y_{k-1} - \mu, \ldots, Y_{k-\lambda+1} - \mu]'$, F is companion matrix

$$F \equiv \begin{bmatrix} B_1 & B_2 & \ldots & B_{\lambda-1} & B_\lambda \\ I & 0 & \ldots & 0 & 0 \\ 0 & I & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & I & 0 \end{bmatrix} \tag{5.32}$$

I is a identity matrix and $V_k \equiv [u_k, 0, \ldots, 0]'$.

If the eigenvalues of F are within the boundary of the unit circle, $F^s \to 0$ as $s \to \infty$, the VAR(1) will be stationary. Under the Wold Theorem, if the VAR(λ) is stationary, then VAR model has an infinite order moving average representation:

$$Y_k = \mu + u_k + \hat{B}_1 u_{k-1} + \hat{B}_2 u_{k-2} \cdots + \hat{B}_\lambda u_{k-\lambda} + \ldots \tag{5.33}$$

$$= \hat{\mathbf{B}}(L)u_k \tag{5.34}$$

where L is the lag operator and $\hat{\mathbf{B}}(L)$ matrix polynomial in L. Using this VMA format, it is convenient for me to figure out the reaction of this system to a shock, namely the impulse response. Assuming the time is $k + s$ now, the VMA system

can be expressed as

$$Y_{k+s} = \mu + u_{k+s} + \hat{B}_1 u_{k+s-1} + \hat{B}_2 u_{k+s-2} + \cdots + \hat{B}_s u_k + \hat{B}_{s+1} u_{k-1} + \dots \quad (5.35)$$

$$= \hat{\mathbf{B}}(L) u_{k+s} \quad (5.36)$$

hence, the VMA format can be equivalently rewritten as follow:

$$Y_k - \mu = \hat{\mathbf{B}}(L) u_k \quad (5.37)$$

where $\hat{\mathbf{B}} \in \{\hat{B}_1, \hat{B}_2, \dots, \hat{B}_s, \hat{B}_{s+1}, \dots\}$, the operators $\hat{\mathbf{B}}(L)$ and $\mathbf{B}(L)$ are related by $\hat{\mathbf{B}}(L) = [\mathbf{B}(L)]^{-1}$. $\hat{B}_s = F_{11}^{(s)}$ is the s order moving average matrix, and $F_{11}^{(s)}$ denotes the upper left block of F^s. Therefore, the $\hat{B}_{ij}^{(s)}$ is the impulse response of $Y_{i,k+s}$ to one unit change in Y_{jk} at the s step.

Following Hasbrouck [1993] pricing error method, the variance of the random-walk component (m_t) can be computed as

$$\sigma_w^2 = [B]_1 Cov(u)[B]_1'], \quad (5.38)$$

where $[B]_1$ is the first row of B matrix. The pricing error variance can be calculated as

$$\sigma_s^2 = \sum_{j=0}^{\infty} [\hat{B}] Cov(u)[\hat{B}]', \quad (5.39)$$

where $\hat{B}_j = -\sum_{k=j+1}^{\infty} [B_j]_1$.

After the calculation of the variance of the pricing error from the vector autoregression, I adopt the Boehmer and Wu [2013] method to produce a proxy of the relative pricing error, $\mathcal{E}$, as the ratio of standard deviation of the pricing error σ_s, to the standard deviation of trade price σ_p, as follow $\mathcal{E} = \sigma_s / \sigma_p$. This pricing

error proxy, $\mathcal{E}$, measures the distance of the real trade prices tracking the efficient prices, hence, the smaller the $\mathcal{E}$, the more efficient the price formation process and the better the market quality.

5.3.4 Multivariate Linear Regression

Maturity effects are an important empirical artefact that must be counted for in any study of a futures market (see Bessembinder et al., 1996; Chen and Zheng, 2013; Kalev and Duong, 2008). Most prior studies focus on the relations between price volatility and time to maturity. This study demonstrates that the efficiency of price discovery varies over the time to maturity. To analyse this effect, I utilize a multivariate linear regression to identify the decomposition maturity effects on price discovery and market quality.

$$\tilde{Y} = c + \tilde{T} + p^{-1} + \mathcal{V} \tag{5.40}$$

The dependent variables ($\tilde{Y}$) include two components: the price discovery indicators and the market quality indicators. The price discovery indicators cover the pricing error ($\mathcal{E}$), the autocorrelation of quote prices (tick-by-tick bids, asks, and mid prices) and the autocorrelation of mid-price return at 26 varying time intervals from tick-by-tick to 30 minutes. The market quality indicators include the number of quotes (n^q) and the number of trades (n) , pricing error ($\mathcal{E}$), quote-to-trade ratio ($\mathcal{Q}$), liquidity spreads (bid-ask spreads, quoted half spreads, quoted depth, effective half spreads, realized spreads, and adverse selection) and the autocorrelation of liquidity spreads. Both realized spreads and adverse selection are also generated at 26 different time intervals. The explanatory variables contain the days to maturity ($\tilde{T}$), the inverse of transaction prices (p^{-1}), and the daily

volatility ($\mathcal{V}$).

All variables are the daily volume-weighted average results across five level order book data. Quote price autocorrelation, mid-quote return autocorrelation, and liquidity spreads autocorrelation are computed at first order, fifth order, and tenth order. The number of trades and quotes, pricing error, the inverse of price, the days to maturity, and the quote-to-trade ratio take the logarithm value.

5.4 Data Integrity and Economic Implications

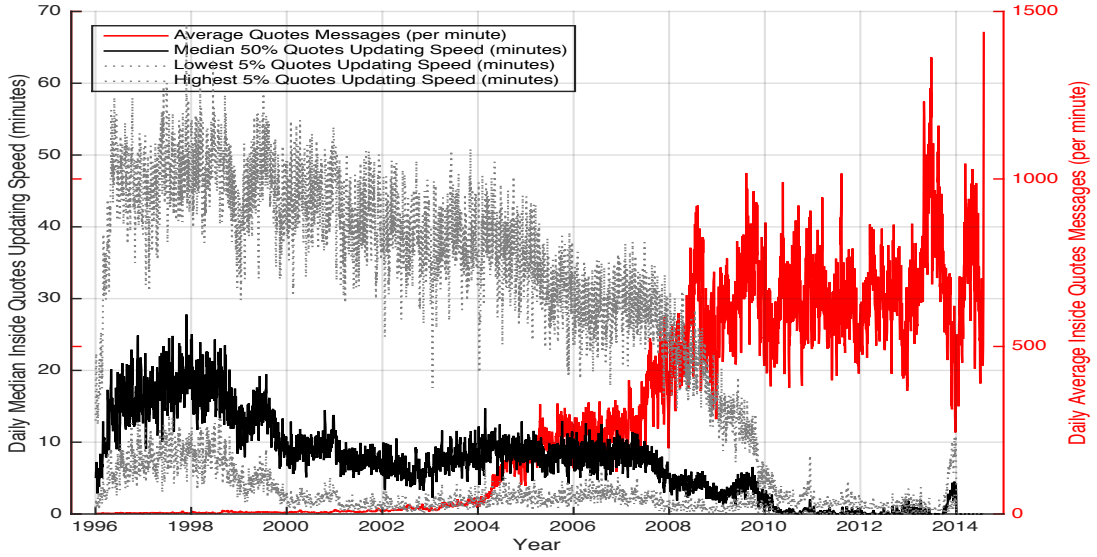
In this section I will introduce the facets of the data that vary from the preceding chapter and then proceed with the interpretation of empirical results. The quoting and trading activity (in terms of numbers of quotes and trades and their individual and aggregate volume) places the data set in the domain of ultra-high frequency analysis. The Eurodollar futures tick data in millisecond timestamps is the same dataset used previously. However, for this exercise, I will not be aggregating the market indicators to the daily level, but instead I will analyze the market data at the tick frequency of the trades. This chapter combines the trade data and limit order book data described in the previous chapters from January 1, 2008 to July 31, 2014.

5.4.1 Data and Summary Statistics

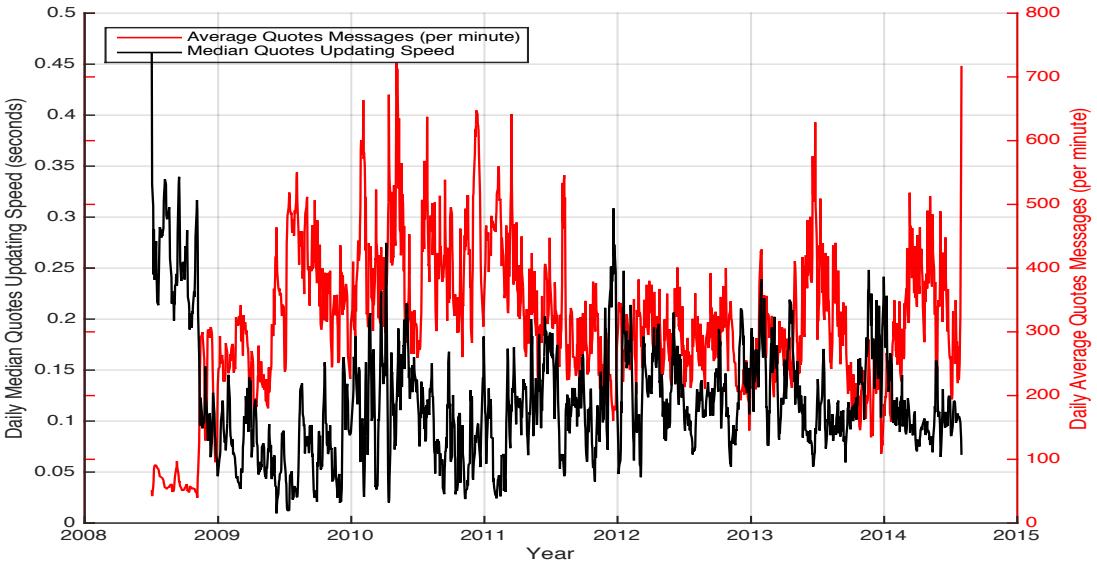
Figure 5.3 presents the daily average quotes messages per minute and the median of updating speeds from 1996 to 2014 for the whole market across the main 40 quarterly dated deliveries. Figure 5.3 Panel (a) depicts the inside quotes activities

and Panel (b) describes the whole limit order book quotes activities from 2008 to 2014. The average number of messages per minute is below 100 messages before 2004, and once it passes 2004, the number of messages within one minute increases dramatically from 115 messages in 2004 to 1,464 messages per minute in 2014. The median of inside quotes updating frequency drops from 19.86 minutes in 1996 to 600 milliseconds in 2014; after 2008 the median of inside quoting frequency maintains a level within 5 minutes and after 2010 the updating speed drops below in 17 seconds. Figure 5.3 Panel (b) represents a faster updating speed based on the whole limit order book from 2008 to 2014. After taking the average across the five levels of the order book, the median of Eurodollar quotes updating speeds are now all below 0.5 seconds. Therefore, Figure 5.3 demonstrates the quarterly settled Eurodollar future are traded with the ultra-high-frequency speeds in milliseconds.

Table 5.2 reports summary statistics for the 40 Eurodollar contracts from 2008 to 2014. The table displays the maturity, the average volume and observation of market depth asks and bids, the average of trades volume, and the average observation of trades for GE contracts having their delivery months in the same year. Specifically, GE?0 includes four 10-year future contracts, namely, GEH0, GEM0, GEU0, and GEZ0 (see Table 5.2). Their delivery months are March 2010, June 2010, September 2010, and December 2010, respectively. The average number of GE?0 bids and asks are both 20.46 million approximately, and the number of trades is 4.10 million. The average volume of GE?0 trades, asks, and bids are \$59.91 trillion, \$58.75 quadrillion, and \$58.58 quadrillion, respectively. Across the 40 GE contracts by year of delivery, GE?4 has the largest bid and ask volume with \$631.86 quadrillion and \$618.16 quadrillion. The largest trades volume is the GE?5 contract with \$75.12 trillion. Meanwhile, GE?5 (all quarters delivering in 2015) contract has the largest average number of bids, ask, and trades, with approximately 57.13 million observations for bids sides and asks



(a) Inside Quotes Activities



(b) Order Book Quotes Activities

Figure 5.3: Quotes Activities

Notes: This figure depicts both the frequency and the speed of Eurodollar future quotes activities from 1996 to 2014. The continuous red line represents the daily average of quotes messages per minute across all ED contracts, the black line is the median of quotes updating speed, and the two grey dotted lines are the 5% and 95% median quotes updating speeds. Panel A illustrates the inside quotes (best-bid and best-ask) messages and its updating speeds (in minutes) from 1996 to 2014. Panel B presents the average of five levels limit order messages and the median updating speed (in seconds) of the whole limit order book from 2008 to 2014.

Table 5.2: Size Data Sample for Limit Order Book and Transaction

		Limit Order Book					
		Asks		**Bids**		**Trades**	
RIC	**Maturity**	Average Volume ($ quadrillion)	Average Number of Obs (million)	Average Volume ($ quadrillion)	Average Number of Obs (million)	Average Volume ($ trillion)	Average Number of Obs (million)
GE?0	2010	58.75	20.46	58.58	20.46	59.91	4.10
GE?1	2011	139.01	32.07	137.53	32.07	66.03	6.42
GE?2	2012	156.87	42.73	155.26	42.73	63.76	8.55
GE?3	2013	326.52	39.07	316.77	39.07	50.01	7.81
GE?4	2014	631.86	47.52	618.16	47.52	62.11	9.50
GE?5	2015	408.38	57.13	409.42	57.13	75.12	11.43
GE?6	2006	149.00	45.21	149.93	45.21	51.36	9.04
GE?7	2007	32.85	22.64	32.88	22.64	19.65	4.53
GE?8	2008	5.51	9.97	5.50	9.97	13.22	2.00
GE?9	2009	7.34	12.56	7.18	12.56	41.22	2.51

Notes: This table depicts data sample size for limit order book and trades data. The roll-over years of the 40 Eurodollar quarterly trade futures and the average number of observations plus the average volume of bids and asks. The limit order book data are from the Thomson Reuters Tick History database from July 1, 2008 to July 31, 2014, labelled by the Reuters Instrument Code (RIC; GEH0, GEH1, ..., GEZ9). The exchange ticker is GE, which is the RIC for the Globex only trades and quotes. I calculate the average value for GE contracts having their maturity date in the same year, denoted GE?0, GE?1, ..., GE?9. Hence, GE?0 includes four 10-year future contracts from 2000 to 2010 – namely, GEH0, GEM0, GEU0, and GEZ0. Note, that the minimum tick size on the exchange is 1/4 of a basis point for the nearest expiring contract and 1/2 otherwise. This yields 32,712 days across 40 contracts for my final daily frequency regression analysis.

sides, and 11.43 million observations of trades. In total, there are 32,712 days with all 40 contracts for the final daily frequency regression analysis.

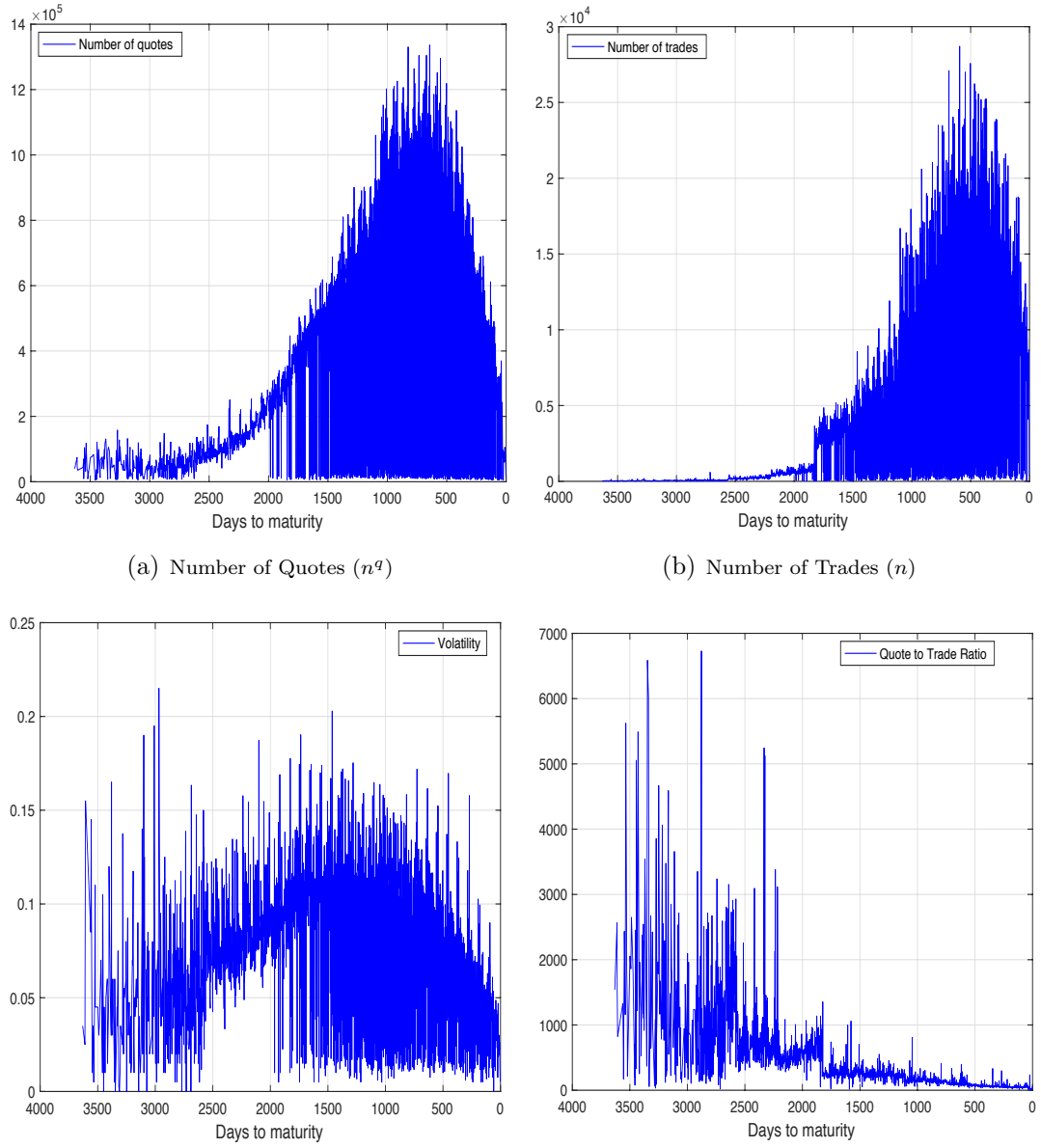
5.4.2 Market Quality and Price Discovery

Figure 5.4 and Table 5.3 report the daily average quotes number, trades number, volatility, and quotes to trades ratio sorted by days to maturity across 40 contracts. Table 5.3 indicates that the average and median of these four measures over different periods, from 30 days before maturity to the whole 10-year sample period. The mean and median of the number of quotes are 569,754 and 474,484 for the entire sample, respectively; the mean and median of the number of trades are 8,107 and 4,195. In Figure 5.4, daily quotes and trades number reach the peak around 1.336 million observations at 644 days to maturity and 28,707 observations 593 days to maturity. This indicates that the highest number of traders submit orders and make transactions around 1.7 – 1.8 years to maturity. It is useful to note the number of quotes is generated by the average or median value across five order book levels. On the 644 days to maturity, the total quotes observation of the full limit order book is around 6.681 million, which translates into nearly 77 quotes entering the order book every second.

Meanwhile, the mean and median of the volatility are 0.09 and 0.08 for the entire sample period. Figure 5.4 indicates that the trade prices volatility has the increasing trend before 1,500 days to maturity, and reaches the highest volatility around three years to two and half years and then drops after two years to maturity. Quote-to-trade ratio describes that the degree of the volume of orders entering the market filled in the transaction executions. The higher value of the quote-to-trade ratio, the lower trade volumes, and the higher quotes volumes. Table

5.3 indicates that the mean and median of the quote-to-trade ratio is 274.83 and 125.35 for the 10-year sample. Quote-to-trade ratio reaches a peak with 6,729.6 on the 2,878 days to maturity, which means the volume of quotes is 6,729.6 times the volume of trades executions (see Figure 5.4). As the contracts approach maturity, the volume of trades rises dramatically. Until 115 days to maturity, the quote-to-trade ratio is 6.7909 around three orders of magnitude lower than the highest quote-to-trade ratio. Table 5.3 also provides the quotes and trades number, volatility, and quotes-to-trades ratio with ten subsamples from 30 days before maturity to 1,800 days before maturity. The number of trade and quotes rises from 30 days to 1,080 days group, and then has a slight decrease at 1,800 days to the maturity group. Both trade price volatility and quote-to-trade ratio monotonously increase from the 30 days subsample to the 1,800 days subsample.

From classical market efficiency analysis when markets are strong form price efficient, previous prices should have no information incorporated in the current prices. To analyze the price efficiency in the Eurodollar future market, I measure the tick-by-tick autocorrelation of bid-price, ask-price , and mid-price with different subsamples from 30 days before the maturity to 1,800 days to maturity. Table 5.3 includes the mean and median of mid-price first-, fifth-, and tenth-order autocorrelations for both the volume-weighted average and each order book level. The volume-weighted average mid-price first-order autocorrelation of the entire sample is 0.9267, which is much lower than it in other lag length subsamples. The mean of VWA mid-price autocorrelation rises from 0.9910 in 30 days group to 0.9945 in 180 days group and then drops to 0.9687 in the 1,800 days group. This translates into the efficiency of price being improved from 30 days to 180 days, but it becomes worse after half year to maturity. The median results point out the similar transformation with a different turning point. The median of VWA mid-price autocorrelation rises from 0.9984 in 30 days group to 0.9995 in 480 days



(a) Number of Quotes (n^q)

(b) Number of Trades (n)

(c) Volatility ($\mathcal{V}$)

(d) Quote-to-trade Ratio ($\mathcal{Q}$)

Figure 5.4: Basic Market Quality Indicators

Notes: This figure represents the daily average quotes number, trades number, volatility, and quotes to trades ratio over 40 Eurodollar contracts sorted by days to maturity. The daily trade price volatility is the difference between the highest transaction price and the lowest transaction price within one day t, $\mathcal{V} = p_t^h - p_t^l$, where p_t^h and p_t^l refer to the highest and lowest trade prices at day t. The quote-to-trade ratio is the number of quotes to the number of trades, $\mathcal{Q} = n_t^q / n_t$, where n_t^q is the daily number of quotes at day t, and n is the daily number of trades at day t.

Table 5.3: Eurodollar Contracts Basic Statistics Description

		Number of quotes n^q	Number of trades n	Volatility $\mathcal{V}$	Quote-to-trade Ratio $\mathcal{Q}$
Mean	**Full samples**	569,753.63	8,106.83	0.09	274.83
	30 days	74,397.26	7,737.58	0.03	18.78
	60 days	177,120.80	8,605.24	0.03	36.26
	90 days	221,844.08	9,280.79	0.04	40.66
	180 days	325,023.59	11,299.10	0.04	44.72
	360 days	473,406.13	13,601.79	0.05	50.32
	480 days	559,660.88	14,729.08	0.06	53.78
	720 days	684,475.94	15,146.39	0.07	67.36
	900 days	737,442.63	14,711.78	0.08	81.94
	1080 days	766,987.31	14,034.30	0.08	95.79
	1800 days	681,650.81	10,266.01	0.09	158.71
Median	**Full samples**	474,483.50	4,194.50	0.08	125.35
	30 days	53,084.00	6,348.50	0.01	7.49
	60 days	106,559.50	6,447.00	0.01	20.22
	90 days	184,772.00	6,983.00	0.02	22.07
	180 days	307,517.50	8,607.00	0.03	25.40
	360 days	422,164.50	10,794.00	0.03	29.57
	480 days	485,629.00	11,591.00	0.04	32.06
	720 days	614,361.50	11,802.50	0.04	38.78
	900 days	684,061.00	11,515.00	0.06	45.32
	1080 days	736,653.00	11,068.00	0.06	52.14
	1800 days	641,881.00	7,022.00	0.08	87.39

Notes: This table demonstrates reports the average and median of quotes number (n^q), trades number (n), volatility ($\mathcal{V}$), and quotes to trades ratio ($\mathcal{Q}$) across 40 Eurodollar contracts. These four measurements have been reported based on the different days to maturity subsamples, from 30 days before maturity to the 10-year Eurodollar contract life cycle. The volatility is the daily highest transaction price subtracting the daily lowest transaction price, $\mathcal{V} = p_t^h - p_t^l$; and the quote-to-trade ratio is the daily total quotes number divided by the daily total trades number, $\mathcal{Q} = n_t^q / n_t$, where p_t^h and p_t^l refer to the highest and lowest trade prices; n_t^q and n_t are the daily number of quotes and transactions at day t, and t is daily time index.

group, and then drops to 0.9983 in the 1,800 days group. A similar trend has been found in each level results and in fifth-order and tenth-order autocorrelations.

Looking at Table 5.3 horizontally, the autocorrelations of mid prices decrease monotonically from level one to level five in all subsamples. However, the results from the whole sample do not follow this rule. Both average and median mid-price autocorrelations decline from level one to level two, then increase at level three and level four and drops to the bottom at level five. For instance, the first-order

Table 5.4: Mid-quoted Price Autocorrelation (at Tick Level)

		Mean						Median					
		VWA	**Level 1**	**level 2**	**Level 3**	**Level 4**	**Level 5**	**VWA**	**Level 1**	**level 2**	**Level 3**	**Level 4**	**Level 5**
First order	**Full samples**	0.9267	0.9348	0.7963	0.8413	0.8886	0.4650	0.9961	0.9937	0.9924	0.9949	0.9961	0.4141
	30 days	0.9910	0.9706	0.9571	0.9289	0.8657	0.5717	0.9984	0.9971	0.9969	0.9965	0.9944	0.6151
	60 days	0.9926	0.9770	0.9658	0.9422	0.9050	0.5839	0.9986	0.9968	0.9967	0.9965	0.9959	0.6578
	90 days	0.9942	0.9810	0.9720	0.9580	0.9266	0.5745	0.9989	0.9971	0.9970	0.9970	0.9965	0.6303
	180 days	0.9945	0.9834	0.9747	0.9668	0.9414	0.5731	0.9992	0.9977	0.9976	0.9976	0.9976	0.6312
	360 days	0.9934	0.9850	0.9805	0.9727	0.9501	0.5734	0.9994	0.9981	0.9980	0.9980	0.9979	0.6190
	480 days	0.9928	0.9842	0.9807	0.9738	0.9536	0.5674	0.9995	0.9982	0.9982	0.9981	0.9980	0.6019
	720 days	0.9899	0.9817	0.9731	0.9650	0.9508	0.5497	0.9994	0.9981	0.9980	0.9979	0.9978	0.5555
	900 days	0.9881	0.9789	0.9623	0.9574	0.9479	0.5265	0.9994	0.9979	0.9978	0.9977	0.9976	0.5138
	1,080 days	0.9867	0.9766	0.9543	0.9506	0.9430	0.5030	0.9993	0.9977	0.9976	0.9975	0.9975	0.4761
	1,800 days	0.9687	0.9690	0.8838	0.8794	0.9033	0.4650	0.9983	0.9962	0.9959	0.9960	0.9964	0.4132
Fifth order	**Full samples**	0.8189	0.8284	0.7048	0.7476	0.8093	0.2900	0.9824	0.9759	0.9712	0.9814	0.9856	0.1419
	30 days	0.9735	0.9427	0.9202	0.8791	0.8046	0.4062	0.9928	0.9880	0.9868	0.9841	0.9773	0.2336
	60 days	0.9783	0.9520	0.9341	0.9066	0.8597	0.4199	0.9940	0.9873	0.9862	0.9858	0.9829	0.3118
	90 days	0.9811	0.9580	0.9450	0.9283	0.8887	0.4092	0.9949	0.9879	0.9877	0.9877	0.9862	0.2963
	180 days	0.9818	0.9621	0.9510	0.9406	0.9106	0.3993	0.9964	0.9902	0.9901	0.9901	0.9895	0.2753
	360 days	0.9782	0.9634	0.9572	0.9480	0.9207	0.3964	0.9972	0.9920	0.9919	0.9917	0.9912	0.2836
	480 days	0.9759	0.9612	0.9561	0.9475	0.9227	0.3915	0.9974	0.9925	0.9924	0.9922	0.9918	0.2692
	720 days	0.9685	0.9552	0.9436	0.9319	0.9125	0.3739	0.9972	0.9923	0.9921	0.9917	0.9912	0.2401
	900 days	0.9635	0.9497	0.9293	0.9193	0.9051	0.3490	0.9969	0.9916	0.9915	0.9912	0.9907	0.2093
	1,080 days	0.9598	0.9454	0.9188	0.9082	0.8954	0.3261	0.9965	0.9910	0.9908	0.9906	0.9902	0.1804
	1,800 days	0.9185	0.9182	0.8320	0.8038	0.8313	0.2903	0.9920	0.9856	0.9846	0.9853	0.9867	0.1417
Tenth order	**Full samples**	0.7560	0.7632	0.6614	0.7111	0.7742	0.2295	0.9666	0.9581	0.9506	0.9691	0.9762	0.0518
	30 days	0.9562	0.9215	0.8997	0.8600	0.7854	0.3479	0.9869	0.9787	0.9751	0.9718	0.9642	0.1360
	60 days	0.9638	0.9321	0.9147	0.8880	0.8400	0.3615	0.9888	0.9784	0.9762	0.9749	0.9701	0.1523
	90 days	0.9685	0.9397	0.9271	0.9109	0.8708	0.3488	0.9903	0.9795	0.9791	0.9790	0.9765	0.1337
	180 days	0.9708	0.9458	0.9349	0.9248	0.8946	0.3378	0.9931	0.9833	0.9829	0.9829	0.9820	0.1167
	360 days	0.9658	0.9468	0.9407	0.9317	0.9037	0.3309	0.9946	0.9864	0.9863	0.9861	0.9852	0.1182
	480 days	0.9620	0.9441	0.9390	0.9305	0.9052	0.3270	0.9949	0.9871	0.9870	0.9868	0.9862	0.1149
	720 days	0.9520	0.9361	0.9241	0.9114	0.8910	0.3105	0.9945	0.9869	0.9866	0.9862	0.9854	0.1030
	900 days	0.9449	0.9293	0.9089	0.8969	0.8815	0.2858	0.9938	0.9859	0.9857	0.9852	0.9845	0.0860
	1,080 days	0.9403	0.9246	0.8983	0.8853	0.8711	0.2644	0.9930	0.9847	0.9845	0.9842	0.9837	0.0699
	1,800 days	0.8830	0.8810	0.8031	0.7715	0.7987	0.2300	0.9843	0.9753	0.9737	0.9757	0.9780	0.0516

Notes: This table reports the mean and median of first-, fifth-, and tenth-order mid-quotes autocorrelations at tick-by-tick time interval. Except for the mid-quotes autocorrelations for the whole sample period, this chapter also compares the autocorrelations within different tenor periods from 30 days to 1,800 days to maturity. Besides, this table also reports the mid-quotes autocorrelations using both the volume-weighted average (VWA) prices and quotes price in each order book level.

autocorrelation of the mid-price in the full sample decreases from 0.2670 at level one to 0.7963 at level two, then increases to 0.8886 at level four, and finally drops to 0.4650 at level five. This indicates that the price in level two is more efficient at absorbing current information and exhibits less influence from the previous price, which indicates the importance of the level two in the limit order book compared with the inside quotes (best bid and best ask). The drop at level five is mainly due to the lack of trading activities and the low quotes updates frequency.

Besides the mid-price autocorrelations, this chapter also explores the bid prices (p_b) and ask prices (p_a) autocorrelations using tick-by-tick time intervals. Figure 5.5 portrays both the volume-weighted average of bid-side and ask-side price autocorrelation and quotes price autocorrelation from level one to level five of the limit order book. Subplots in the first row are the volume-weighted average of quotes price first-order, fifth-order, and tenth-order autocorrelations. The blue line represents the bid-side autocorrelation, and the red line is the ask prices autocorrelation, while the second row presents the quotes price autocorrelations in each order book level. Each subplot displays a butterfly pattern in where the right wing represents bid-side prices autocorrelations, and the left wing represents ask-side prices autocorrelations. For both sides, quotes autocorrelations are increasing as the tenor decreases.

More specifically, the highest ask price and bid price first-order autocorrelations are 0.9997 on around one year before maturity (364 days and 376 days before maturity, respectively). Figure 5.5 also presents negative autocorrelations in each subplot. For the first-order volume-weighted average autocorrelation, the lowest ask-side and bid-side price autocorrelations are -0.3601 on 3,022 days to maturity and -0.2474 on 3,304 days to maturity. The one with negative autocorrelation with the nearest day to maturity are -0.0569 on 2,857 days for ask-side and -0.0696

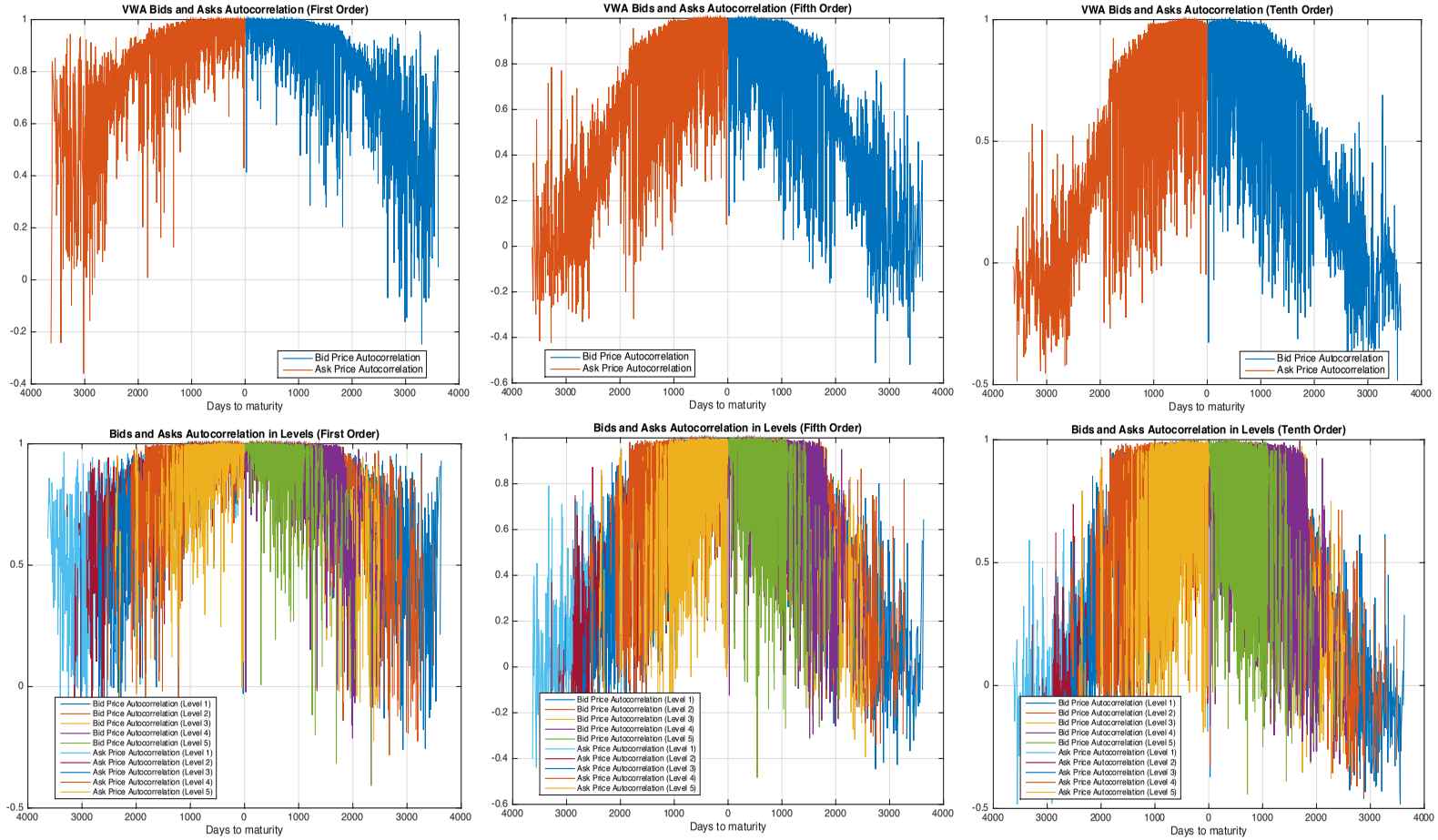


Figure 5.5: Bid and Ask Prices Autocorrelations

Notes: This plot portrays the volume weighted average (VWA) asks and bids autocorrelations and the bids and asks autocorrelations for each level in the limit order book. The bid and ask autocorrelations are calculated at tick-by-tick basis firstly, and then take the daily average value for each contracts each trading day. Then I sort all the daily average autocorrelations by the days to maturity, and take the average value across all 40 Eurodollar contracts. The first row plots are the first-order, fifth-order, and tenth-order autocorrelations of the bid-side (blue line) and ask-side (red line) volume weighted average prices. The second row plots are the first-order, fifth-order, and tenth-order autocorrelations of bid and ask prices from level 1 to level 5 limit order book. The left side represents ask-price autocorrelations from level one to level five, and the right side represents bid-price autocorrelations for each level.

on 2,667 days for bid-side. The fifth-order and tenth-order price autocorrelations have lower and more negative autocorrelation compared with first-order autocorrelations. Ask-price and bid-price autocorrelations in levels (see the second row in Figure 5.5) show the similar pattern with the volume-weighted average ones. The closer to the maturity date, the higher price autocorrelations are for both sides. Besides, the fifth-order and tenth-order autocorrelations are lower than at first-order autocorrelations for each level.

Prior studies point out that the return autocorrelation is a core indicator as a proxy for price efficiency [Anderson et al., 2013; Foley and Putniņš, 2014; Hendershott and Jones, 2005]. A relatively small return autocorrelation (less than 0.2) indicates the higher efficiency of the prices and the better price formation process in the market; whereas, the higher autocorrelation is, the lower efficiency of the future prices. Hendershott and Jones [2005] report the first-order autocorrelation of mid-quote returns over a quarter-minute, half-minute and one-minute intervals from the NBBO of the US ETFs. Their findings suggest that the quote return autocorrelation at 30 seconds interval is between 0.04 and 0.19. Foley and Putniņš [2014] illustrate that the mean and median of absolute autocorrelations in the mid-price returns span between 0.04 and 0.06 at a ten-second time interval. In this chapter, I calculate the first-order, fifth-year and tenth-order autocorrelations of mid-quote returns at varying timescales from tick-by-tick to half hour. I also measure the return autocorrelations in different subsamples (such as 30-day, 60-day, 90-day, . . . , 1,800-day before maturity). To avoid unnecessary replication, here, I only report the full sample return autocorrelations. However, the results are consistent across the several subsamples and available in the online appendix.

Table 5.5 summarizes both the average and median of mid-quote return autocorrelations utilizing the volume weighted average mid prices and mid prices of

each level in the limit order book. Figure 5.6 also portrays the change of the volume weighted average return autocorrelations across different time intervals. On a tick-by-tick timescale, both the volume-weighted average mid-quote return and mid-quote return in each order book level are all negatively correlated with the previous return. The average of first-order mid-price return autocorrelation is -0.0526 at tick-by-tick time scale, which is consistent with prior studies (see Bianco and Renò, 2006; Lo and MacKinlay, 1988). The return autocorrelation across five levels also are negative. This indicates the previous first-order return decreasingly impacts the current return.

The negative mid-price return autocorrelation also suggests that, if the previous return is above the average level, the current return is more likely to be below the average. In a high-frequency market, this is a common situation on short-horizon returns, implying a fast speed mean-reversion of the bid-ask price at the tick level in the Eurodollar market. Furthermore, the price is efficient in terms of reflecting available information in the market due to the small absolute value of autocorrelation. I can conclude from this structure that the deviations from the efficient price are persistent and significant, but short lived. However, there is a noticeable term structure of this persistence.

Recall, that the use of the ED contract in constructing synthetic swaps results in clustering of necessary transactions over the term structure (at regular intervals by IMM date and at regular tenors, such as one, three and five years, it appears that this increases in liquidity draws in trading that is both faster and results in more significant deviations from efficiency, but the deviations decay rapidly. We can see that the autocorrelations reported herein are substantially higher than those reported in the equity markets by [Anderson et al., 2013; Foley and Putniņš, 2014; Hendershott and Jones, 2005], by an order of between 3 and 5

times; a range of ≈ 0.3 to ≈ 0.9 for time frames above tick versus 0.04 to 0.19 in the equity literature.

Except for the results at tick-by-tick, the others are all positively autocorrelated. From the shortest time interval (1 ms) to the longest time interval (1,800 sec), the autocorrelations are positive and gradually increase. More specifically, Table 5.5 describes that the first order autocorrelations of VWA mid-price returns increase from 0.3549 (1 ms interval) to 0.8925 (1,800 sec interval). In Figure 5.6, the return autocorrelations rise monotonously at the tick-by-tick level. The autocorrelation increases rapidly from 1 ms to 2 ms. It experiences a slower increase from 5 ms to 10 ms, and then increases steadily until 1,800-second timescale (see Figure 5.6). Negative tick by tick autocorrelations are fascinating, but consider that in the most active periods the negative autocorrelation will last for less than 1 to 10ms on average, which is very quick.

The mid-price return the autocorrelations decrease from level one to level two, subsequently increase from level three to level four and finally drop at level five. Table 5.5 and Figure 5.5 provide very similar patterns in the first-order, fifth-order, and tenth-order autocorrelations. Both the red line (fifth-order autocorrelations) and black line (tenth-order autocorrelations) are below the first-order return autocorrelations, which suggests that the fifth-order and tenth-order autocorrelations are more efficient of the price adjustment. However, the median fifth-order and tenth-order autocorrelations in Figure 5.6 are lower than the mean autocorrelations from 1 ms to 10 ms.

Figure 5.7 presents the daily volume-weighted average of mid-quote return autocorrelation across 40 Eurodollar contracts in six selected time intervals (tick-by-tick, 25 ms, 200 ms, 1 sec, 60 sec, 1,800 sec). The mid-quote return autocorrelation plots with other timescales are available in the online appendix. We can

see the highest mid-quote return autocorrelation in each subplot occurs at the first-order autocorrelation within one year to the maturity period. Since the time interval increases, the mid-quote return autocorrelation also increases, which is consistent with Table 5.5. More specifically, the tick level return autocorrelations have a large amount of negative results, especially in the first-order autocorrelations. When contracts approach the maturity days, the autocorrelation at three different orders seems to converge to zero, which implies higher efficiency of the price to incorporate available information in the market.

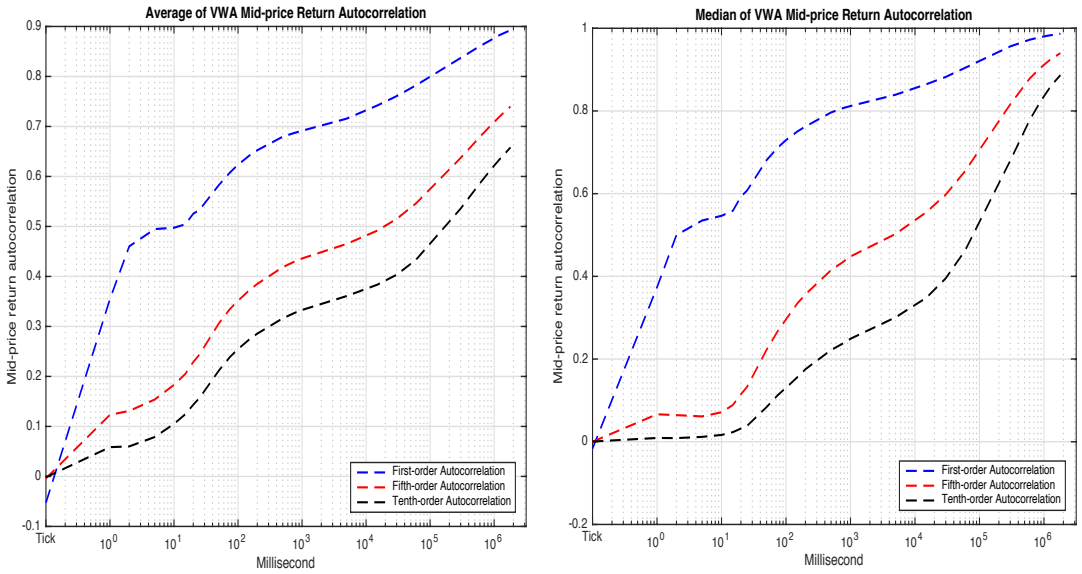


Figure 5.6: Volume Weighted Average Mid-quoted Return Autocorrelation (Mean and Median of All 40 Eurodollar Contracts)

Notes: This figure portrays both the average and median of first-order, fifth-order, and tenth-order volume weighted average mid-quote return autocorrelations for the full sample. The volume weighted average mid-price is calculated by the sum of $p_{mjk}v_{mjk}$ and weighted by the daily total quotes volume for the day, where k is the intraday time index and j is the order book level. This plot represents the average and median of return autocorrelations over 40 Eurodollar contracts. The first-order, fifth-order, and tenth-order mid-quote return autocorrelations are marked with blue, red, and black dashed lines. Each order mid-quote return autocorrelations are measured at 26 time intervals, including tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec.

Table 5.5: Mid-quoted Return Autocorrelation (Full sample)

		Mean						Median					
		VWA	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**	**VWA**	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
	Tick	-0.0526	-0.0889	-0.1425	-0.1470	-0.1242	-0.3651	-0.0163	-0.0919	-0.1076	-0.1132	-0.1014	-0.4167
	1ms	0.3549	0.3221	0.3210	0.3271	0.3337	0.2457	0.3737	0.3334	0.3333	0.3340	0.3627	0.1996
	2ms	0.4602	0.3864	0.3825	0.3957	0.4085	0.2611	0.4998	0.4375	0.4349	0.4496	0.4708	0.2691
	5ms	0.4947	0.4479	0.4371	0.4634	0.4843	0.2843	0.5351	0.4994	0.4988	0.5001	0.5327	0.3143
	10ms	0.4974	0.4826	0.4699	0.5045	0.5299	0.3383	0.5464	0.5268	0.5134	0.5576	0.5871	0.3633
	15ms	0.5042	0.4960	0.4801	0.5212	0.5514	0.3643	0.5587	0.5430	0.5287	0.5744	0.6123	0.3889
	20ms	0.5254	0.5176	0.5005	0.5444	0.5780	0.3887	0.5928	0.5714	0.5602	0.6100	0.6492	0.4166
	25ms	0.5331	0.5260	0.5089	0.5557	0.5910	0.4099	0.6079	0.5866	0.5719	0.6284	0.6679	0.4358
	50ms	0.5825	0.5797	0.5586	0.6132	0.6538	0.4747	0.6811	0.6670	0.6577	0.7194	0.7579	0.4996
	75ms	0.6080	0.6018	0.5805	0.6398	0.6839	0.5044	0.7119	0.6946	0.6808	0.7540	0.7939	0.5283
	100ms	0.6234	0.6153	0.5931	0.6560	0.7028	0.5205	0.7296	0.7082	0.6938	0.7728	0.8120	0.5463
	150ms	0.6417	0.6328	0.6094	0.6746	0.7230	0.5370	0.7504	0.7236	0.7115	0.7924	0.8329	0.5678
First	**200ms**	0.6524	0.6434	0.6189	0.6864	0.7338	0.5473	0.7624	0.7323	0.7217	0.8030	0.8445	0.5847
Order	**500ms**	0.6795	0.6684	0.6404	0.7117	0.7600	0.5751	0.7964	0.7533	0.7437	0.8275	0.8669	0.6253
	750ms	0.6869	0.6760	0.6474	0.7185	0.7680	0.5867	0.8065	0.7604	0.7509	0.8332	0.8735	0.6384
	1s	0.6914	0.6805	0.6504	0.7218	0.7716	0.5938	0.8119	0.7638	0.7539	0.8363	0.8764	0.6476
	5s	0.7157	0.7019	0.6657	0.7378	0.7876	0.6253	0.8393	0.7873	0.7813	0.8503	0.8902	0.6768
	15s	0.7420	0.7296	0.6868	0.7542	0.8037	0.6464	0.8646	0.8185	0.8119	0.8685	0.9019	0.6959
	30s	0.7607	0.7505	0.7020	0.7680	0.8163	0.6567	0.8823	0.8426	0.8356	0.8854	0.9136	0.7027
	60s	0.7821	0.7737	0.7205	0.7840	0.8300	0.6625	0.9044	0.8691	0.8625	0.9052	0.9265	0.7061
	300s	0.8370	0.8369	0.7669	0.8184	0.8604	0.6666	0.9561	0.9368	0.9315	0.9509	0.9601	0.7037
	600s	0.8606	0.8635	0.7860	0.8323	0.8725	0.6661	0.9723	0.9581	0.9543	0.9660	0.9718	0.7008
	900s	0.8732	0.8774	0.7950	0.8363	0.8770	0.6650	0.9789	0.9673	0.9643	0.9728	0.9773	0.6998
	1200s	0.8824	0.8862	0.8009	0.8407	0.8805	0.6669	0.9827	0.9726	0.9698	0.9768	0.9806	0.7023
	1500s	0.8880	0.8924	0.8051	0.8429	0.8829	0.6660	0.9849	0.9760	0.9734	0.9794	0.9828	0.7010
	1800s	0.8925	0.8970	0.8083	0.8448	0.8857	0.6645	0.9867	0.9784	0.9760	0.9814	0.9846	0.7004
	Tick	-0.0039	-0.0142	-0.0131	-0.0154	-0.0173	-0.0059	0.0001	-0.0144	-0.0095	-0.0123	-0.0144	0.0000
	1ms	0.1233	0.0586	0.0581	0.0605	0.0626	0.0088	0.0664	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	2ms	0.1310	0.0940	0.0920	0.0983	0.1038	0.0218	0.0644	0.0000	0.0000	0.0000	0.0000	-0.0001
	5ms	0.1538	0.1405	0.1374	0.1525	0.1653	0.0585	0.0613	0.0250	0.0188	0.0540	0.0799	0.0000
	10ms	0.1832	0.1751	0.1708	0.1947	0.2146	0.0993	0.0713	0.0647	0.0567	0.1028	0.1366	0.0000
	15ms	0.2049	0.1957	0.1905	0.2204	0.2446	0.1205	0.0889	0.0896	0.0795	0.1348	0.1760	0.0229
	20ms	0.2288	0.2180	0.2127	0.2463	0.2744	0.1396	0.1149	0.1234	0.1112	0.1726	0.2227	0.0499
	25ms	0.2437	0.2299	0.2240	0.2615	0.2922	0.1505	0.1332	0.1382	0.1247	0.1955	0.2504	0.0666
	50ms	0.3058	0.2829	0.2757	0.3242	0.3633	0.1882	0.2225	0.2367	0.2193	0.3149	0.3747	0.1213
	75ms	0.3350	0.3038	0.2961	0.3511	0.3949	0.2049	0.2684	0.2707	0.2499	0.3641	0.4341	0.1478
	100ms	0.3514	0.3159	0.3078	0.3670	0.4138	0.2157	0.2960	0.2864	0.2666	0.3919	0.4655	0.1630
	150ms	0.3724	0.3324	0.3241	0.3880	0.4378	0.2319	0.3345	0.3055	0.2877	0.4258	0.5051	0.1882
Fifth	**200ms**	0.3851	0.3433	0.3345	0.4016	0.4529	0.2422	0.3565	0.3210	0.3022	0.4471	0.5286	0.2047
Order	**500ms**	0.4185	0.3723	0.3623	0.4359	0.4915	0.2732	0.4150	0.3591	0.3405	0.4943	0.5811	0.2532
	750ms	0.4293	0.3823	0.3722	0.4477	0.5053	0.2868	0.4338	0.3730	0.3575	0.5096	0.5978	0.2696
	1s	0.4360	0.3876	0.3771	0.4536	0.5119	0.2962	0.4481	0.3774	0.3642	0.5186	0.6048	0.2848
	5s	0.4651	0.4154	0.4020	0.4822	0.5436	0.3428	0.5026	0.4154	0.4016	0.5501	0.6400	0.3499
	15s	0.4917	0.4464	0.4282	0.5090	0.5725	0.3773	0.5552	0.4720	0.4576	0.5862	0.6700	0.4015
	30s	0.5151	0.4742	0.4517	0.5339	0.5968	0.4011	0.5977	0.5212	0.5061	0.6280	0.7010	0.4301
	60s	0.5459	0.5101	0.4836	0.5652	0.6284	0.4216	0.6547	0.5843	0.5700	0.6815	0.7424	0.4543
	300s	0.6370	0.6212	0.5779	0.6495	0.7104	0.4583	0.8160	0.7745	0.7611	0.8250	0.8535	0.4824
	600s	0.6792	0.6709	0.6181	0.6834	0.7436	0.4675	0.8785	0.8452	0.8341	0.8766	0.8961	0.4852
	900s	0.7029	0.6985	0.6387	0.6984	0.7582	0.4711	0.9054	0.8771	0.8680	0.9007	0.9156	0.4852
	1200s	0.7190	0.7161	0.6523	0.7086	0.7682	0.4767	0.9219	0.8959	0.8869	0.9146	0.9276	0.4925
	1500s	0.7307	0.7297	0.6621	0.7153	0.7753	0.4774	0.9321	0.9087	0.8996	0.9241	0.9356	0.4922
	1800s	0.7399	0.7396	0.6698	0.7203	0.7810	0.4790	0.9399	0.9178	0.9098	0.9311	0.9422	0.4917
	Tick	-0.0016	-0.0063	-0.0056	-0.0059	-0.0064	-0.0018	0.0005	-0.0048	-0.0009	-0.0014	-0.0026	0.0000
	1ms	0.0586	0.0220	0.0217	0.0229	0.0239	-0.0009	0.0092	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	2ms	0.0600	0.0363	0.0357	0.0387	0.0413	0.0023	0.0088	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	5ms	0.0788	0.0677	0.0664	0.0745	0.0822	0.0192	0.0118	-0.0001	-0.0001	0.0000	0.0000	-0.0001
	10ms	0.1051	0.0960	0.0942	0.1088	0.1218	0.0450	0.0165	0.0000	0.0000	0.0027	0.0228	-0.0001
	15ms	0.1243	0.1135	0.1112	0.1307	0.1476	0.0606	0.0233	0.0000	0.0000	0.0253	0.0506	0.0000
	20ms	0.1434	0.1304	0.1279	0.1510	0.1710	0.0730	0.0320	0.0174	0.0135	0.0474	0.0804	0.0000
	25ms	0.1568	0.1404	0.1378	0.1641	0.1865	0.0804	0.0387	0.0286	0.0234	0.0649	0.1037	0.0000
	50ms	0.2115	0.1833	0.1803	0.2169	0.2469	0.1074	0.0838	0.0868	0.0772	0.1520	0.2053	0.0191
	75ms	0.2390	0.2018	0.1988	0.2414	0.2757	0.1185	0.1132	0.1109	0.0999	0.1912	0.2524	0.0319
	100ms	0.2547	0.2121	0.2092	0.2552	0.2925	0.1255	0.1305	0.1244	0.1114	0.2148	0.2831	0.0400
	150ms	0.2740	0.2257	0.2227	0.2735	0.3141	0.1373	0.1561	0.1410	0.1288	0.2440	0.3200	0.0559
Tenth	**200ms**	0.2856	0.2343	0.2313	0.2848	0.3271	0.1441	0.1754	0.1529	0.1390	0.2608	0.3409	0.0661
Order	**500ms**	0.3162	0.2587	0.2557	0.3171	0.3646	0.1680	0.2219	0.1830	0.1695	0.3071	0.3969	0.1016
	750ms	0.3266	0.2681	0.2652	0.3291	0.3785	0.1784	0.2369	0.1949	0.1819	0.3218	0.4164	0.1192
	1s	0.3329	0.2729	0.2699	0.3352	0.3855	0.1852	0.2491	0.1991	0.1876	0.3305	0.4243	0.1307
	5s	0.3609	0.2988	0.2942	0.3649	0.4197	0.2233	0.3009	0.2327	0.2195	0.3648	0.4660	0.1885
	15s	0.3835	0.3261	0.3187	0.3915	0.4494	0.2534	0.3483	0.2817	0.2664	0.4066	0.5043	0.2357
	30s	0.4034	0.3510	0.3411	0.4167	0.4762	0.2765	0.3953	0.3300	0.3135	0.4540	0.5439	0.2698
	60s	0.4332	0.3864	0.3734	0.4513	0.5128	0.2988	0.4629	0.3951	0.3786	0.5172	0.5977	0.2981
	300s	0.5357	0.5090	0.4823	0.5608	0.6238	0.3484	0.6798	0.6345	0.6168	0.7178	0.7611	0.3432
	600s	0.5864	0.5686	0.5330	0.6076	0.6705	0.3629	0.7780	0.7407	0.7251	0.7967	0.8280	0.3485
	900s	0.6146	0.6010	0.5598	0.6296	0.6926	0.3692	0.8251	0.7921	0.7772	0.8350	0.8600	0.3501
	1200s	0.6334	0.6227	0.5771	0.6446	0.7072	0.3767	0.8537	0.8229	0.8082	0.8578	0.8796	0.3572
	1500s	0.6468	0.6382	0.5892	0.6541	0.7173	0.3777	0.8728	0.8439	0.8301	0.8733	0.8924	0.3585
	1800s	0.6581	0.6509	0.5993	0.6616	0.7255	0.3808	0.8866	0.8582	0.8452	0.8850	0.9031	0.3582

Notes: This table reports the mean and median of first-order, fifth-order, and tenth-order autocorrelations of mid-quote return for the 10-year full sample period across all 40 Eurodollar contracts, using both the volume weighted average (VWA) mid-price (p^v_{mk}) and mid-quote prices (p_{mjk}) in each level of limit order book, where k is the intraday time index and j is the order book level. For each order, mid-quote return autocorrelations are measured at varying time intervals, including tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec. This table only reports the full sample return autocorrelations; however, the results are consistent across the diverse subsamples and available in the online appendix, including 30 days to 1,800 days before maturity.

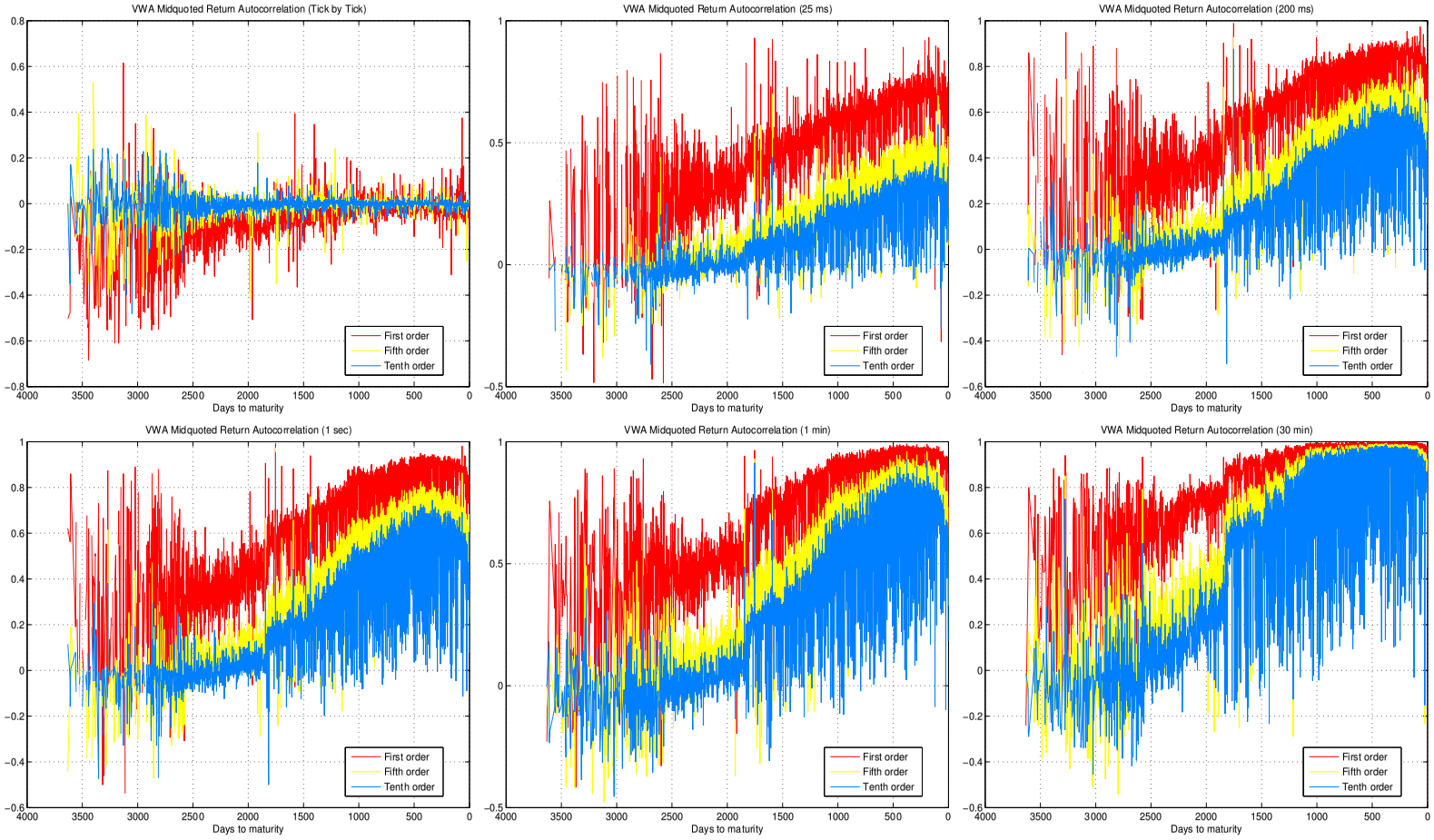


Figure 5.7: Daily Volume Weighted Average Mid-quoted Return Autocorrelation (Sorted by Days to Maturity)

Notes: This figure represents the daily of the volume weighted average (VWA) of mid-quote return autocorrelations sorted by the days to maturity across 40 Eurodollar contracts. The first-order, fifth-order, and tenth-order mid-quote return autocorrelations are marked with red, yellow, and blue lines. The first step is to utilize the volume weighted average mid-price (p^v_{mk}) to get the mid-price returns ($r^v_{\tau k}$) at 26 varying time intervals $\tau \in \{$tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec$\}$. The next step is to calculate the volume weighted average return autocorrelations, $\Omega^{R,v}_{\tau\eta k} = \rho(r^v_{\tau k}, r^v_{\tau k-\eta})$, where k is the intraday time index and η is the orders of autocorrelations $\eta \in \{1^{st}, 5^{th}, 10^{th}\}$. Then, I take the average return autocorrelations for each day and each contract, and obtain 32,856 daily return autocorrelations based on the actual trading days for 40 contracts. Finally, I sort the 32,856 daily return autocorrelations by the days to maturity $\tilde{T}$, and obtain 2,434 average daily return autocorrelations over all 40 contracts. Because of the 10-year life cycle of Eurodollar contracts, the days to maturity spans from 2 days before maturity to time span 3,631 days to maturity. The figure only reports the volume weighted average return autocorrelations at six selected time intervals (tick-by-tick, 25 ms, 200 ms, 1 sec, 60 sec, 1,800 sec). The mid-quote return autocorrelations plots with other timescales are available in the online appendix.

To have a comprehensive understanding of the market liquidity of the Eurodollar market, this chapter presents a series of liquidity spreads as metrics across the complete order book, covering the bid-ask spreads ($\tilde{\mathscr{S}}^{Q}$), quoted half-spread ($\tilde{\mathscr{S}}^{Q^{1/2}}$), quoted depth ($\tilde{\mathscr{S}}^{D}$), effective half-spread ($\tilde{\mathscr{S}}^{E}$), realized spreads ($\tilde{\mathscr{S}}^{R}$), and adverse selection ($\tilde{\mathscr{S}}^{AS}$). Both realized spreads and adverse selection are measured at diverse timescales from tick level to 1,800 seconds. I also consider the first-order, fifth-order, and tenth-order autocorrelations of these various liquidity spreads. Both liquidity spreads and their autocorrelation have also been computed with different subsamples, such as 30-day, 60-day, 90-day, 180-day, 360-day, 480-day, 720-day, 900-day, 1,080-day, and 1,800-day before maturity. These liquidity measurements are available in the online appendix. Here I only report the full sample liquidity spreads and its autocorrelations in different order distance. Table 5.6 displays the average and median of the liquidity spreads between volume-weighted average across five levels and in each level. All spreads are recorded in basis points except for the quoted depth.

In Table 5.6, the median of volume-weighted average bid-ask spreads is 190.3462 bps, while the median of bid-ask spread for each order book level is much narrower, such as 0.5662 bps in level one and 1.5954 bps in level two. Both the average and median of bid-ask spreads have a widening tendency, moving from level one to level five. A similar pattern has been found in the quoted half spreads. This suggests that the inside quotes of Eurodollar market are more liquid than other levels. However, I find contrary evidence from the quoted depth. Both the mean and median of quoted depth in level two are much larger than in the inside quotes; the quoted depth in level two is the largest one compared with other levels. Specifically, the depth of inside quotes is 694,049, and the average of level two quoted depth is 1,461,332 which is more than twice higher than the depth of level one. Moreover, the median of level two depth is approximately 3.6 times

larger than the median of level one quoted depth; even the median of level three depth is higher than the first level. This implies that level two has the highest ability to absorb large orders before moving the prices and provides the largest liquidity in the Eurodollar market, even better than the best bids and best asks, which again emphasizes the importance of the level two.

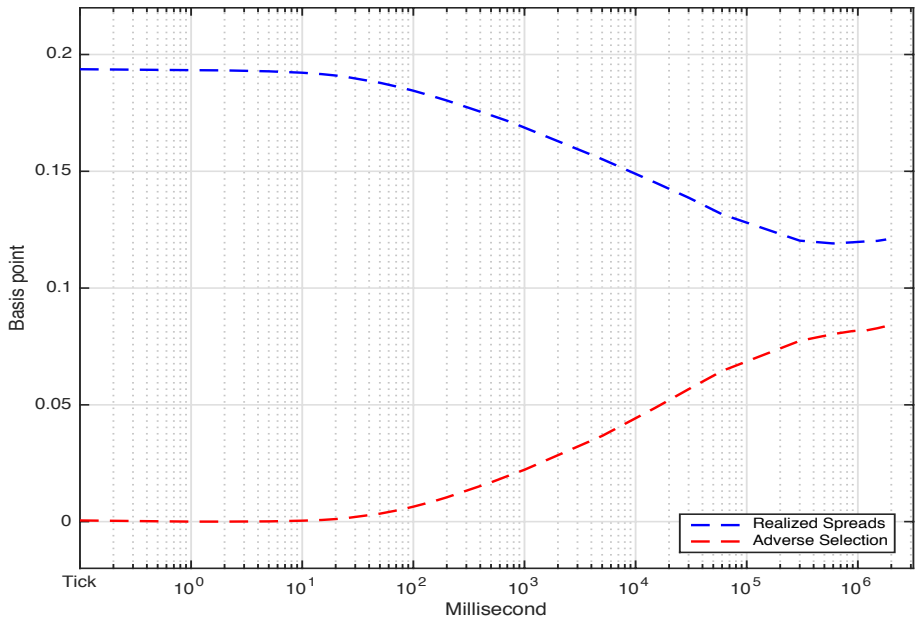


Figure 5.8: Median Volume Weighted Average Market Liquidity Spreads

Notes: This figure displays the median of the volume weighted average realized spreads (blue dashed line) and adverse selection (red dashed line) at varying time intervals of all 40 Eurodollar contracts. Both realized spreads and adverse selection are measured with the volume weighted average bid, ask, and mid-quote (p^v_{bk}, p^v_{ak}, and p^v_{mk}) with trade prices (p_k), where k is the intraday time index (see Equation 5.3). The directions of trades are classified by the volume weighted average trade classification algorithm (see Equation 5.4). The 26 time intervals includes tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec.

The median of volume-weighted average effective half spreads is 0.1933 bps. The median of level two effective half spreads with 0.1950 bps is narrower than the median of effective half spread in level one with 0.1970 bps. The smaller effective half spread, the more liquid is the order book level, and the more actual transac-

Table 5.6: Market Liquidity Spreads (Full sample)

		Mean						Median					
		VWA	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**	**VWA**	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
$\tilde{\mathscr{S}}^{Q}$		614.5158	0.8968	2.2715	6.3150	12.2411	20.8223	190.3462	0.5662	1.5954	2.6317	3.6476	4.6516
$\tilde{\mathscr{S}}^{Q^{1/2}}$		307.2579	0.4484	1.1358	3.1575	6.1206	10.4111	95.1731	0.2831	0.7977	1.3158	1.8238	2.3258
$\tilde{\mathscr{S}}^{D}$		8,013,956,100	694,049	1,461,332	371,477	113,273	98,955	779,834	62,689	226,449	94,131	26,917	24,212
$\tilde{\mathscr{S}}^{E}$		2.1138	0.1028	-7.7191	-48.4440	-128.4050	-138.5255	0.1933	0.1970	0.1950	0.1976	0.1989	-4.4198
	Tick	2.1297	0.1061	-6.8225	-40.3641	-107.9635	-117.5563	0.1937	0.1860	0.1840	0.1873	0.1892	0.0212
	1ms	2.1138	0.1028	-7.7074	-48.4674	-128.3911	-138.4424	0.1933	0.1970	0.1950	0.1975	0.1988	-4.2821
	2ms	2.1138	0.1026	-7.6627	-48.5399	-128.0210	-138.2548	0.1932	0.1968	0.1949	0.1974	0.1986	-4.1845
	5ms	2.1130	0.1021	-7.6292	-48.4518	-127.8014	-137.6092	0.1928	0.1963	0.1944	0.1968	0.1981	-3.6087
	10ms	2.1113	0.1012	-7.5836	-48.6501	-127.2178	-138.3436	0.1922	0.1953	0.1935	0.1956	0.1971	-2.8894
	15ms	2.1056	0.1003	-7.5812	-48.7468	-126.9956	-137.5359	0.1916	0.1943	0.1925	0.1944	0.1958	-2.3996
	20ms	2.1042	0.0998	-7.5745	-48.7852	-127.1197	-136.8907	0.1910	0.1933	0.1913	0.1932	0.1946	-2.0594
	25ms	2.0973	0.0985	-7.5261	-48.8403	-127.2109	-137.1840	0.1904	0.1921	0.1901	0.1919	0.1933	-1.6595
	50ms	2.0867	0.0930	-7.4065	-48.9752	-126.5558	-137.4107	0.1879	0.1864	0.1845	0.1854	0.1864	-0.7530
	75ms	2.0844	0.0888	-7.3033	-48.7910	-127.4522	-135.4844	0.1860	0.1818	0.1798	0.1802	0.1809	-0.5739
	100ms	2.0771	0.0853	-7.6281	-48.7263	-126.0619	-133.9055	0.1845	0.1782	0.1760	0.1760	0.1767	-0.3931
	150ms	2.0737	0.0810	-7.7250	-48.2544	-126.0175	-134.9254	0.1821	0.1726	0.1703	0.1696	0.1702	-0.2131
$\tilde{\mathscr{S}}^{R}$	**200ms**	2.0724	0.0768	-7.3706	-48.2108	-123.6547	-134.0619	0.1803	0.1687	0.1662	0.1646	0.1653	-0.1967
	500ms	2.1095	0.0695	-7.5536	-48.2980	-121.5839	-132.6012	0.1740	0.1566	0.1531	0.1492	0.1498	-0.0171
	750ms	2.1128	0.0559	-7.7341	-47.6592	-121.6615	-132.0049	0.1711	0.1513	0.1475	0.1430	0.1435	0.0441
	1sec	2.0994	0.0476	-7.5538	-47.5738	-122.7269	-130.8667	0.1687	0.1477	0.1437	0.1387	0.1392	0.0552
	5sec	2.0788	0.0623	-7.6832	-46.9353	-123.5844	-127.5781	0.1552	0.1285	0.1229	0.1153	0.1155	0.0815
	15sec	2.0523	0.1053	-7.2485	-45.5941	-120.6437	-127.0673	0.1452	0.1150	0.1093	0.0999	0.1000	0.0710
	30sec	2.0090	0.1134	-6.0448	-47.7588	-118.7421	-125.3470	0.1387	0.1065	0.1011	0.0916	0.0918	0.0622
	60sec	1.9385	0.0979	-5.5505	-46.3371	-115.6659	-130.4785	0.1316	0.0992	0.0943	0.0850	0.0857	0.0624
	300sec	1.9214	0.0890	-4.9691	-42.2134	-101.7786	-117.1717	0.1203	0.0906	0.0872	0.0814	0.0813	0.0658
	600sec	2.0202	0.1187	-6.6919	-40.1396	-92.4790	-116.7235	0.1191	0.0913	0.0881	0.0823	0.0834	0.0685
	900sec	1.8071	0.0045	-6.2287	-34.4263	-87.4582	-103.9156	0.1196	0.0933	0.0904	0.0861	0.0854	0.0743
	1200sec	1.8017	0.0236	-6.9650	-38.9689	-91.8655	-107.9092	0.1200	0.0948	0.0919	0.0879	0.0859	0.0721
	1500sec	1.7773	0.0171	-6.7447	-41.2462	-83.0098	-102.0246	0.1202	0.0955	0.0928	0.0877	0.0873	0.0756
	1800sec	1.7545	0.0139	-5.7499	-43.1279	-82.1438	-106.0423	0.1208	0.0962	0.0940	0.0894	0.0884	0.0753
	Tick	0.0161	0.0054	0.8621	9.9047	18.1384	24.0045	0.0005	-0.0061	-0.0058	-0.0049	-0.0042	1.3749
	1ms	0.0000	0.0001	-0.0116	0.0455	0.0175	-0.0830	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2ms	0.0001	0.0002	-0.0556	0.1498	-0.0166	-0.1824	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	5ms	0.0008	0.0007	-0.0901	0.1371	0.0242	-0.6923	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000
	10ms	0.0025	0.0016	-0.1339	0.0639	-0.1337	-1.3636	0.0004	0.0006	0.0006	0.0006	0.0006	0.0000
	15ms	0.0082	0.0025	-0.1132	0.3058	-0.1803	-1.8222	0.0007	0.0014	0.0014	0.0014	0.0014	0.0000
	20ms	0.0096	0.0030	-0.1223	0.2602	-0.3743	-2.2685	0.0011	0.0020	0.0020	0.0021	0.0021	0.0000
	25ms	0.0165	0.0043	-0.1730	0.3595	-0.4117	-2.4063	0.0015	0.0029	0.0030	0.0031	0.0031	0.0000
	50ms	0.0271	0.0096	-0.2955	0.6153	-0.5136	-3.9982	0.0034	0.0065	0.0069	0.0074	0.0075	0.0000
	75ms	0.0294	0.0138	-0.4459	0.5283	-0.2929	-4.4485	0.0050	0.0094	0.0099	0.0109	0.0110	0.0000
	100ms	0.0367	0.0173	-0.1393	-0.2015	-0.4199	-5.0892	0.0064	0.0117	0.0124	0.0137	0.0137	0.0000
	150ms	0.0401	0.0212	0.0811	-0.6045	-0.0851	-5.9282	0.0086	0.0151	0.0160	0.0178	0.0177	0.0000
$\tilde{\mathscr{S}}^{AS}$	**200ms**	0.0415	0.0252	0.0729	-0.7424	-0.5690	-5.4917	0.0104	0.0177	0.0188	0.0211	0.0208	0.0000
	500ms	0.0043	0.0307	0.1920	-0.9007	-1.5959	-6.9593	0.0168	0.0271	0.0289	0.0327	0.0320	0.0000
	750ms	0.0010	0.0450	0.2815	-1.0461	-1.9166	-8.4584	0.0200	0.0316	0.0338	0.0381	0.0371	0.0000
	1sec	0.0144	0.0527	0.1183	-1.2677	-1.5976	-8.4537	0.0222	0.0349	0.0371	0.0422	0.0408	0.0000
	5sec	0.0351	0.0348	-0.0017	-2.3230	-3.4511	-13.9229	0.0367	0.0540	0.0578	0.0658	0.0635	0.0000
	15sec	0.0615	-0.0053	-0.2574	-3.9900	-4.8729	-12.0516	0.0487	0.0692	0.0730	0.0818	0.0798	0.0000
	30sec	0.1048	-0.0123	-1.4300	-4.7015	-7.6117	-16.3323	0.0567	0.0793	0.0826	0.0919	0.0892	0.0000
	60sec	0.1753	0.0021	-1.7156	-5.1425	-10.3501	-15.7121	0.0645	0.0883	0.0912	0.0997	0.0966	0.0000
	300sec	0.1924	0.0127	-2.6812	-9.0430	-21.2597	-19.0821	0.0774	0.1033	0.1041	0.1078	0.1051	0.0000
	600sec	0.0937	-0.0173	-2.0077	-10.8883	-23.0613	-26.3751	0.0803	0.1046	0.1051	0.1067	0.1038	0.0000
	900sec	0.3067	0.0982	-1.6569	-13.4521	-26.4347	-31.6533	0.0816	0.1040	0.1037	0.1052	0.1025	-0.0154
	1200sec	0.3122	0.0796	-1.5112	-12.1490	-25.1776	-34.9664	0.0819	0.1044	0.1038	0.1044	0.1023	-0.0134
	1500sec	0.3365	0.0875	-1.7325	-11.1023	-25.0916	-31.4382	0.0828	0.1046	0.1035	0.1043	0.1012	-0.0419
	1800sec	0.3593	0.0917	-2.4525	-10.5097	-28.2102	-34.7940	0.0837	0.1040	0.1024	0.1037	0.0998	-0.0937

Notes: This table describes that the average and median of a range of market liquidity spreads within the full sample period, including bid-ask spreads ($\tilde{\mathscr{S}}^{Q}$), quoted half-spread ($\tilde{\mathscr{S}}^{Q^{1/2}}$), quoted depth ($\tilde{\mathscr{S}}^{D}$), effective half-spread ($\tilde{\mathscr{S}}^{E}$), realized spreads ($\tilde{\mathscr{S}}^{R}$), and adverse selection ($\tilde{\mathscr{S}}^{AS}$). All the spreads are calculated by the volume weighted average (VWA) quote prices and the prices for order book level 1 to level 5. Both realized spreads and adverse selection are measured at 26 varying time intervals: tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec. Except for the quotes depth, all spreads are reported in basis point (*bps*). Liquidity spreads for other tenor subsamples, such as 30 days to maturity and 1,800 days to maturity, are available in the online appendix.

tion prices are near the mid price. Figure 5.8 depicts the changes in the median of the volume weighted average realized spreads and adverse selection at varying time-horizon. The median of realized spread (blue dashed line) displays a decreased tendency, and the median of adverse selection (red dashed line) increases monotonously from the short-horizon to long-term timescale. When the time interval is below 25 milliseconds, both the realized spreads and adverse selection keep steady. Once the timescale passes 25 ms, the median of volume-weighted average realized spreads drops from 0.1879 bps at 50 ms to 0.1196 bps at 900 seconds (see Table 5.6). The median of adverse selection rises from 0.0034 bps at 50 ms to 0.0816 bps at 900 seconds. Both realized spreads and adverse selection become stable after the timescales are above 5 minutes or 10 minutes.

After measuring the diverse market liquidity spreads, I also compare the spreads autocorrelation with different order distance. The autocorrelation of spreads can reflect the spread changes over different timescales, and the spreads are calculated by the quote price and transaction price. This means that the spread autocorrelation also can reflect the efficiency of the price incorporating the available information in the market. Table 5.7 to Table 5.9 illustrate the first-order, fifth-order, and tenth-order autocorrelations of bid-ask spreads, quoted half spread, quoted depth, effective half spreads, realized spreads at 26 time intervals and adverse selection at 26 time intervals.

Both the average and median of the first-order autocorrelation of the bid-ask spread, quoted half spread, quoted depth, and effective half spread are highly positively correlated above 0.6 at tick level, while the fifth-order and tenth-order autocorrelations of these four spreads are much lower, especially the tenth-order autocorrelations are all below 0.5. The larger positive spreads autocorrelations reflect the previous spread at certain timescale increasingly affects the current

spreads, or the previous spreads is above the average level, the current spreads are more likely to be above the average. In other words, the higher spreads autocorrelation indicates the ability of price absorbing the available information is low.

The three plots in the first row of Figure 5.9 also describe the volume-weighted average of bid-ask spreads, quoted depth, and effective spread autocorrelations by the contract maturity dates. I find the consistent evidence that the fifth-order and tenth-order autocorrelations are below the first-order at tick level, and the autocorrelations rise when the contracts approach the maturity date. Since the date to maturity is less than around 500 days, the volume-weighted average of bid-ask spread autocorrelations at all three different lag distance are highly positive correlated and more than 0.8. In Table 5.7, the first-order bid-ask spreads autocorrelations increase from level one with 0.7749 to level four with 0.6163, and then drop to 0.5353 in level five. In Table 5.8, the most efficient bid-ask spreads adjustment is the median tenth-order autocorrelation with 0.1116 in level one. The other three spreads autocorrelation also follow this pattern.

I also compare the realized spread autocorrelation and adverse selection autocorrelations with three different order distances in Table 5.7 to Table 5.9. Both the first-order realized spread and adverse selection autocorrelations have an increasing tendency from the tick-by-tick timescale to 1,800 seconds. For each level autocorrelations, both spreads autocorrelations rise from level one to level four and decline in level five. The median autocorrelations of both realized spread and adverse selection at the tick-by-tick level are much lower than other time intervals, which indicates the adjustment of spreads and the incorporation of information are much efficient at tick level. One point to note is that the median of tick-by-tick adverse selection autocorrelation for each level is negatively small or

around zero, which implies that the previous adverse selection is less correlated or uncorrelated with the future adverse selection. In other words, the speed of price adjustment to absorb information is at tick level.

The second and third rows of Figure 5.9 display the volume-weighted average of realized spreads and adverse selection autocorrelation by the contract maturity dates at three orders. Here I select three time-interval (tick-by-tick, 200 milliseconds, and one minute); the autocorrelation with other timescales can be found in the online appendix. The first-order autocorrelations are higher than the autocorrelations in the other two. This means the better spread adjustment over time is the longer order distance. Furthermore, the realized spreads and adverse selection autocorrelations are higher approaching the maturity dates (see Figure 5.9). When the contract is close to maturity, the more active traders and more information are in the Eurodollar market. The adjustments of spreads are becoming less sensitive in the face of increasing of available information in the market, which could be the reason for the increasing spread autocorrelation.

To capture how and in what manner the six factors affect each other, I also provide the measurements of the impulse response functions for the trade return, signed trade direction indicator, signed trade volume, signed square root of trade volume, signed ask-side market concentration ratio, and the signed bid-side market concentration ratio respectively (see Figure 5.10 to Figure 5.11). I run the first-order vector autoregression model with 1,000 steps on tick-by-tick basis data for the full sample across 10-year life cycle of Eurodollar contract with 177.5442 million observations. Additionally, I also report the impulse response for the days to maturity within the first year (43.5116 million observations), days to maturity within the third year (42.5787 million observations), and days to maturity within the fifth year (10.4092 million observations), denoted as one-year tenor with the

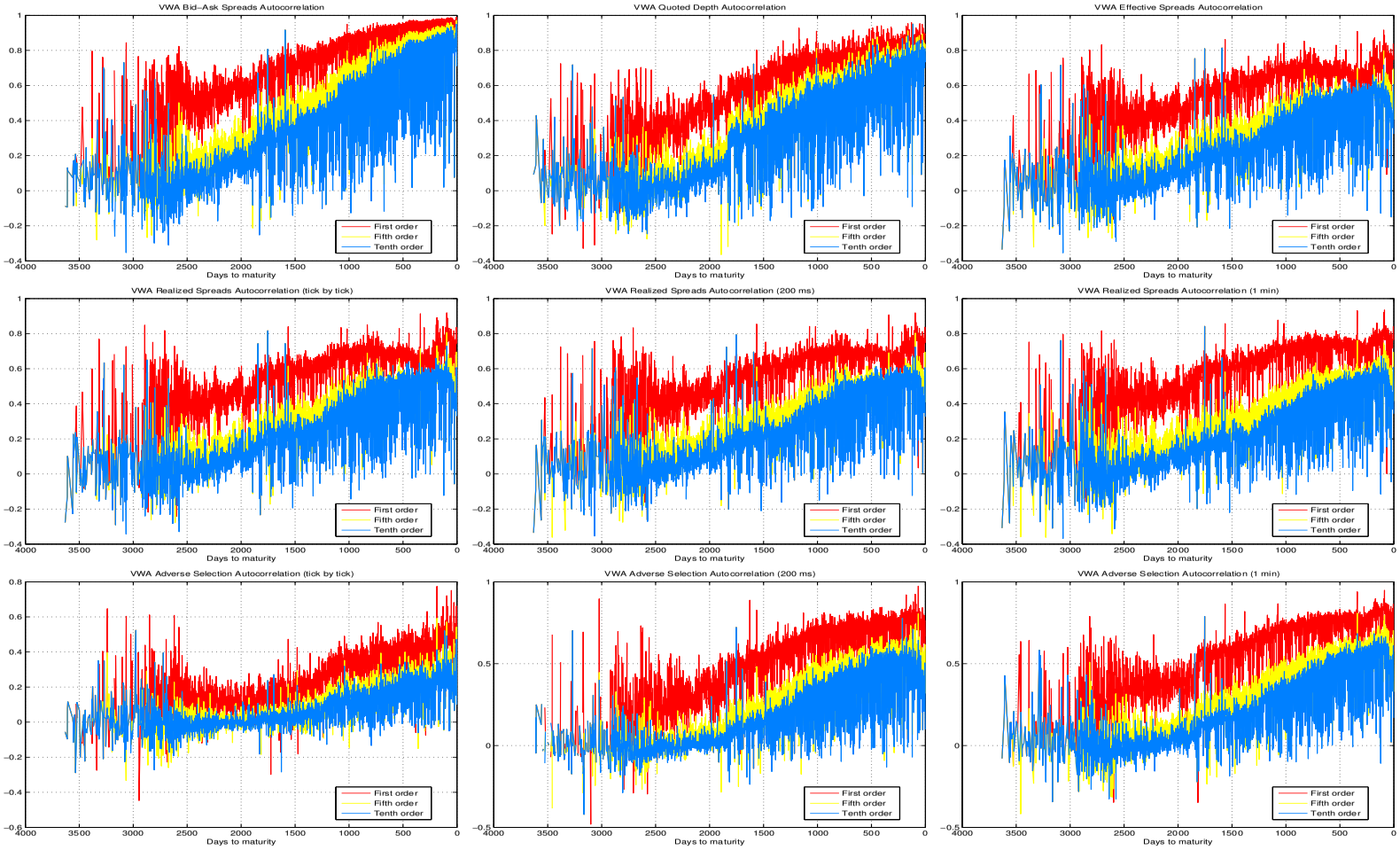


Figure 5.9: Daily Volume-weighted Average Spreads Autocorrelation (Sorted by Days to Maturity)

Notes: This plot represents the daily volume weighted average (VWA) spreads autocorrelations sorted by the days to maturity across 40 Eurodollar contracts, including the first-order, fifth-order, and tenth-order autocorrelations of bid-ask spreads ($\tilde{\mathscr{S}}^{Q}$), quoted half-spread ($\tilde{\mathscr{S}}^{Q^{1/2}}$), quoted depth ($\tilde{\mathscr{S}}^{D}$), effective half-spread ($\tilde{\mathscr{S}}^{E}$), realized spreads ($\tilde{\mathscr{S}}^{R}$), and adverse selection ($\tilde{\mathscr{S}}^{AS}$). The first-order, fifth-order, and tenth-order spreads autocorrelations are marked with red, yellow, and blue lines. The volume weighted average spreads, $\tilde{\mathscr{S}}^{v}_{k}$, are are measured with the volume weighted average bid, ask, and mid-quote (p^{v}_{bk}, p^{v}_{ak}, and p^{v}_{mk}) with trade prices (p_k), where k is the intraday time index (see Equation 5.3). The trade directions are calculated by the volume weighted average trade classification algorithm (see Equation 5.4). Hence, the autocorrelations of the volume weighted average spreads can be expressed as $\Omega^{\tilde{\mathscr{S}},v}_{\eta k} = \rho(\tilde{\mathscr{S}}^{v}_{k}, \tilde{\mathscr{S}}^{v}_{k-\eta})$, where η indicates the first-order, fifth-order, and tenth-order. Because the spreads autocorrelations are taken the average value over 40 Eurodollar contracts sorted by the days to maturity, there are 2,434 daily average spreads autocorrelations in total and the time spans of two days and 3,631 days to maturity. This plot only reports realized spread and adverse selection autocorrelations at three time-interval: tick-by-tick, 200 milliseconds, and one minute. The other spreads autocorrelations are available in the online appendix.

Table 5.7: The First-order Spreads Autocorrelation

		Mean						Median					
		VWA	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**	**VWA**	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
$\Omega^{\tilde{\mathscr{P}}^Q}$		0.7749	0.4985	0.5012	0.5900	0.6163	0.5353	0.8568	0.5100	0.5132	0.6248	0.6517	0.5162
$\Omega^{\tilde{\mathscr{P}}^{Q^{1/2}}}$		0.7749	0.4985	0.5012	0.5900	0.6163	0.5353	0.8568	0.5100	0.5132	0.6248	0.6517	0.5162
$\Omega^{\tilde{\mathscr{P}}^D}$		0.6540	0.7108	0.8464	0.8080	0.9310	0.8654	0.7724	0.7729	0.9349	0.8487	0.9654	0.9001
$\Omega^{\tilde{\mathscr{P}}^E}$		0.6117	0.5556	0.5067	0.5951	0.6455	0.3549	0.6619	0.5808	0.5645	0.6713	0.7217	0.3333
	Tick	0.6073	0.4131	0.4007	0.5369	0.5835	0.4123	0.3970	0.3991	0.5996	0.6507	0.4146	0.0000
	1ms	0.6117	0.5558	0.5069	0.5953	0.6456	0.3522	0.6619	0.5809	0.5646	0.6715	0.7220	0.3308
	2ms	0.6116	0.5560	0.5071	0.5957	0.6461	0.3483	0.6618	0.5812	0.5647	0.6717	0.7225	0.3224
	5ms	0.6114	0.5568	0.5084	0.5967	0.6476	0.3453	0.6616	0.5817	0.5658	0.6736	0.7245	0.3117
	10ms	0.6113	0.5582	0.5109	0.5984	0.6500	0.3486	0.6613	0.5834	0.5693	0.6765	0.7270	0.3130
	15ms	0.6113	0.5597	0.5132	0.6005	0.6526	0.3538	0.6609	0.5851	0.5718	0.6796	0.7301	0.3184
	20ms	0.6113	0.5611	0.5151	0.6021	0.6541	0.3562	0.6610	0.5868	0.5740	0.6827	0.7328	0.3222
	25ms	0.6112	0.5627	0.5173	0.6041	0.6565	0.3619	0.6610	0.5894	0.5766	0.6855	0.7351	0.3306
	50ms	0.6111	0.5696	0.5263	0.6117	0.6651	0.3750	0.6610	0.5968	0.5864	0.6973	0.7471	0.3538
	75ms	0.6113	0.5742	0.5318	0.6167	0.6704	0.3816	0.6612	0.6018	0.5930	0.7061	0.7536	0.3688
	100ms	0.6114	0.5775	0.5361	0.6202	0.6738	0.3872	0.6613	0.6061	0.5974	0.7108	0.7587	0.3786
	150ms	0.6119	0.5820	0.5416	0.6249	0.6789	0.3969	0.6618	0.6114	0.6026	0.7177	0.7665	0.3975
$\Omega^{\tilde{\mathscr{P}}^R}$	**200ms**	0.6122	0.5850	0.5451	0.6284	0.6821	0.4042	0.6620	0.6148	0.6075	0.7216	0.7711	0.4075
	500ms	0.6145	0.5932	0.5553	0.6374	0.6906	0.4328	0.6654	0.6251	0.6201	0.7327	0.7825	0.4503
	750ms	0.6158	0.5960	0.5589	0.6404	0.6935	0.4480	0.6669	0.6285	0.6243	0.7362	0.7858	0.4708
	1sec	0.6174	0.5979	0.5612	0.6421	0.6949	0.4600	0.6694	0.6319	0.6273	0.7392	0.7875	0.4842
	5sec	0.6282	0.6073	0.5744	0.6508	0.7022	0.5327	0.6795	0.6444	0.6426	0.7452	0.7924	0.5680
	15sec	0.6348	0.6119	0.5811	0.6553	0.7053	0.5786	0.6852	0.6496	0.6477	0.7471	0.7928	0.6155
	30sec	0.6376	0.6147	0.5857	0.6574	0.7058	0.6043	0.6875	0.6526	0.6511	0.7476	0.7929	0.6442
	60sec	0.6402	0.6175	0.5901	0.6584	0.7064	0.6250	0.6904	0.6574	0.6557	0.7474	0.7920	0.6666
	300sec	0.6462	0.6288	0.6033	0.6657	0.7102	0.6553	0.6950	0.6718	0.6682	0.7527	0.7925	0.7009
	600sec	0.6490	0.6360	0.6117	0.6712	0.7145	0.6611	0.6976	0.6806	0.6764	0.7572	0.7942	0.7097
	900sec	0.6501	0.6389	0.6145	0.6741	0.7174	0.6628	0.6989	0.6849	0.6797	0.7598	0.7949	0.7141
	1200sec	0.6504	0.6403	0.6179	0.6766	0.7180	0.6664	0.6985	0.6863	0.6816	0.7604	0.7958	0.7177
	1500sec	0.6507	0.6425	0.6217	0.6785	0.7200	0.6669	0.6983	0.6873	0.6839	0.7621	0.7963	0.7155
	1800sec	0.6512	0.6439	0.6234	0.6805	0.7210	0.6666	0.6984	0.6891	0.6845	0.7628	0.7977	0.7168
	Tick	0.2766	-0.0322	-0.0560	-0.0606	-0.0458	-0.1186	-0.0256	-0.0285	-0.0292	-0.0252	-0.1239	0.0000
	1ms	0.3714	0.3007	0.2998	0.3055	0.3117	0.2255	0.4245	0.3324	0.3324	0.3329	0.3332	0.1249
	2ms	0.4319	0.3529	0.3493	0.3617	0.3735	0.2362	0.5013	0.3993	0.3991	0.4091	0.4280	0.2302
	5ms	0.4408	0.4074	0.3985	0.4227	0.4417	0.2559	0.4930	0.4664	0.4592	0.4888	0.4994	0.2681
	10ms	0.4441	0.4431	0.4330	0.4658	0.4903	0.2972	0.4928	0.4973	0.4935	0.5161	0.5453	0.3052
	15ms	0.4539	0.4590	0.4460	0.4854	0.5152	0.3166	0.5083	0.4997	0.4970	0.5393	0.5777	0.3210
	20ms	0.4746	0.4810	0.4670	0.5091	0.5420	0.3347	0.5371	0.5346	0.5206	0.5776	0.6143	0.3374
	25ms	0.4839	0.4901	0.4757	0.5211	0.5560	0.3502	0.5546	0.5502	0.5360	0.5949	0.6378	0.3471
	50ms	0.5288	0.5389	0.5202	0.5754	0.6155	0.3984	0.6206	0.6307	0.6181	0.6839	0.7231	0.3935
	75ms	0.5491	0.5572	0.5385	0.5986	0.6416	0.4198	0.6440	0.6506	0.6395	0.7141	0.7540	0.4178
	100ms	0.5606	0.5686	0.5491	0.6121	0.6583	0.4308	0.6554	0.6578	0.6477	0.7296	0.7712	0.4323
	150ms	0.5737	0.5825	0.5622	0.6279	0.6757	0.4417	0.6672	0.6679	0.6585	0.7452	0.7885	0.4511
$\Omega^{\tilde{\mathscr{P}}^{AS}}$	**200ms**	0.5819	0.5912	0.5703	0.6378	0.6849	0.4482	0.6754	0.6743	0.6643	0.7517	0.7985	0.4624
	500ms	0.6000	0.6083	0.5846	0.6565	0.7053	0.4622	0.6918	0.6849	0.6738	0.7699	0.8142	0.4865
	750ms	0.6044	0.6117	0.5880	0.6602	0.7097	0.4672	0.6952	0.6861	0.6768	0.7719	0.8167	0.4914
	1sec	0.6068	0.6133	0.5885	0.6609	0.7108	0.4699	0.6972	0.6857	0.6762	0.7720	0.8174	0.4943
	5sec	0.6132	0.6136	0.5854	0.6598	0.7104	0.4749	0.6905	0.6772	0.6705	0.7612	0.8094	0.4982
	15sec	0.6181	0.6180	0.5858	0.6584	0.7078	0.4750	0.6813	0.6739	0.6655	0.7542	0.8031	0.4948
	30sec	0.6212	0.6217	0.5862	0.6587	0.7073	0.4740	0.6799	0.6708	0.6630	0.7502	0.7997	0.4939
	60sec	0.6251	0.6245	0.5867	0.6592	0.7058	0.4704	0.6779	0.6712	0.6624	0.7473	0.7949	0.4892
	300sec	0.6351	0.6349	0.5893	0.6599	0.7034	0.4601	0.6834	0.6809	0.6673	0.7485	0.7924	0.4757
	600sec	0.6408	0.6413	0.5929	0.6626	0.7038	0.4571	0.6871	0.6873	0.6718	0.7478	0.7927	0.4752
	900sec	0.6428	0.6433	0.5927	0.6613	0.7035	0.4553	0.6892	0.6908	0.6721	0.7491	0.7931	0.4724
	1200sec	0.6448	0.6441	0.5926	0.6614	0.7038	0.4562	0.6890	0.6913	0.6722	0.7475	0.7919	0.4717
	1500sec	0.6460	0.6457	0.5935	0.6615	0.7043	0.4554	0.6900	0.6919	0.6730	0.7481	0.7929	0.4710
	1800sec	0.6469	0.6465	0.5940	0.6616	0.7056	0.4546	0.6911	0.6929	0.6744	0.7491	0.7941	0.4722

Notes: This table reports both the mean and median of a series of liquidity spreads first-order autocorrelations for the full 10-year sample period, including the first-order autocorrelations of bid-ask spreads ($\Omega^{\tilde{\mathscr{P}}^Q}$), quoted half-spread ($\Omega^{\tilde{\mathscr{P}}^{Q^{1/2}}}$), quoted depth ($\Omega^{\tilde{\mathscr{P}}^D}$), effective half-spread ($\Omega^{\tilde{\mathscr{P}}^E}$), realized spreads ($\Omega^{\tilde{\mathscr{P}}^R}$), and adverse selection ($\Omega^{\tilde{\mathscr{P}}^{AS}}$). All the spreads autocorrelations are calculated by the volume weighted average (VWA) spreads and the spreads for each order book level. Both realized spreads autocorrelations and adverse selection autocorrelations are measured at 26 varying time intervals, including tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec. The first-order spreads autocorrelations for other tenor subsamples (such as 30 days to maturity and 1,800 days to maturity) are available in the online appendix.

Table 5.8: The Fifth-order Spreads Autocorrelation

		Mean						Median					
		VWA	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**	**VWA**	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
$\Omega^{\tilde{\mathscr{F}}Q}$		0.5597	0.2340	0.2486	0.3315	0.3597	0.3285	0.6259	0.2017	0.2149	0.3108	0.3397	0.2249
$\Omega^{\tilde{\mathscr{F}}Q^{1/2}}$		0.5597	0.2340	0.2486	0.3315	0.3597	0.3285	0.6259	0.2017	0.2149	0.3108	0.3397	0.2249
$\Omega^{\tilde{\mathscr{F}}D}$		0.4766	0.4929	0.7164	0.6015	0.8342	0.6948	0.4859	0.5225	0.8467	0.6353	0.8989	0.7557
$\Omega^{\tilde{\mathscr{F}}E}$		0.3789	0.2932	0.2733	0.3464	0.3970	0.1635	0.3955	0.2575	0.2321	0.3565	0.4281	0.0901
$\Omega^{\tilde{\mathscr{F}}R}$	**Tick**	0.3763	0.2435	0.2349	0.3175	0.3654	0.1887	0.1761	0.1598	0.3050	0.3782	0.1013	0.0000
	1ms	0.3789	0.2934	0.2736	0.3467	0.3972	0.1629	0.3955	0.2576	0.2321	0.3568	0.4287	0.0879
	2ms	0.3788	0.2936	0.2738	0.3470	0.3977	0.1630	0.3954	0.2579	0.2323	0.3574	0.4294	0.0864
	5ms	0.3785	0.2942	0.2746	0.3477	0.3989	0.1640	0.3952	0.2585	0.2337	0.3582	0.4317	0.0857
	10ms	0.3781	0.2954	0.2761	0.3491	0.4008	0.1659	0.3947	0.2599	0.2356	0.3604	0.4352	0.0890
	15ms	0.3778	0.2966	0.2776	0.3507	0.4027	0.1666	0.3940	0.2610	0.2376	0.3624	0.4371	0.0900
	20ms	0.3776	0.2980	0.2793	0.3524	0.4045	0.1671	0.3935	0.2621	0.2394	0.3643	0.4400	0.0935
	25ms	0.3774	0.2991	0.2807	0.3539	0.4061	0.1681	0.3935	0.2630	0.2406	0.3672	0.4424	0.0959
	50ms	0.3766	0.3055	0.2878	0.3613	0.4151	0.1723	0.3914	0.2694	0.2472	0.3794	0.4566	0.1090
	75ms	0.3755	0.3098	0.2921	0.3663	0.4209	0.1749	0.3902	0.2728	0.2500	0.3880	0.4674	0.1175
	100ms	0.3749	0.3126	0.2953	0.3699	0.4249	0.1776	0.3889	0.2758	0.2522	0.3929	0.4743	0.1238
	150ms	0.3741	0.3167	0.3000	0.3750	0.4309	0.1830	0.3871	0.2793	0.2561	0.4016	0.4852	0.1345
	200ms	0.3736	0.3197	0.3032	0.3788	0.4354	0.1875	0.3867	0.2814	0.2590	0.4073	0.4924	0.1418
	500ms	0.3728	0.3281	0.3126	0.3895	0.4471	0.2061	0.3852	0.2880	0.2697	0.4215	0.5127	0.1669
	750ms	0.3727	0.3313	0.3161	0.3932	0.4513	0.2162	0.3841	0.2914	0.2743	0.4254	0.5188	0.1803
	1sec	0.3733	0.3333	0.3181	0.3954	0.4537	0.2245	0.3847	0.2944	0.2758	0.4286	0.5222	0.1902
	5sec	0.3791	0.3426	0.3288	0.4066	0.4650	0.2754	0.3868	0.3059	0.2922	0.4379	0.5319	0.2538
	15sec	0.3804	0.3478	0.3350	0.4129	0.4712	0.3125	0.3856	0.3126	0.3004	0.4436	0.5364	0.3002
	30sec	0.3788	0.3506	0.3381	0.4159	0.4732	0.3359	0.3800	0.3178	0.3065	0.4477	0.5393	0.3274
	60sec	0.3782	0.3537	0.3418	0.4177	0.4751	0.3538	0.3791	0.3259	0.3156	0.4499	0.5400	0.3500
	300sec	0.3807	0.3644	0.3518	0.4256	0.4809	0.3793	0.3820	0.3518	0.3397	0.4634	0.5475	0.3798
	600sec	0.3832	0.3716	0.3591	0.4312	0.4849	0.3848	0.3843	0.3649	0.3534	0.4720	0.5505	0.3893
	900sec	0.3853	0.3756	0.3620	0.4332	0.4871	0.3865	0.3872	0.3713	0.3588	0.4758	0.5544	0.3905
	1200sec	0.3862	0.3778	0.3648	0.4353	0.4882	0.3900	0.3899	0.3760	0.3628	0.4782	0.5557	0.3942
	1500sec	0.3871	0.3804	0.3683	0.4363	0.4887	0.3887	0.3908	0.3799	0.3691	0.4790	0.5570	0.3940
	1800sec	0.3882	0.3823	0.3701	0.4376	0.4891	0.3893	0.3920	0.3824	0.3720	0.4809	0.5570	0.3936
$\Omega^{\tilde{\mathscr{F}}AS}$	**Tick**	0.1321	-0.0018	-0.0012	0.0005	0.0008	0.0011	0.0010	-0.0001	0.0000	0.0000	-0.0001	0.0000
	1ms	0.1268	0.0524	0.0519	0.0540	0.0558	0.0108	0.0717	-0.0002	-0.0002	-0.0001	-0.0001	-0.0001
	2ms	0.1167	0.0850	0.0834	0.0892	0.0940	0.0227	0.0518	-0.0001	-0.0001	-0.0001	0.0000	-0.0001
	5ms	0.1290	0.1275	0.1249	0.1387	0.1503	0.0516	0.0381	0.0000	0.0000	0.0379	0.0611	-0.0001
	10ms	0.1560	0.1598	0.1563	0.1786	0.1974	0.0827	0.0453	0.0492	0.0416	0.0835	0.1161	0.0000
	15ms	0.1769	0.1797	0.1753	0.2034	0.2264	0.0998	0.0598	0.0723	0.0626	0.1164	0.1558	0.0068
	20ms	0.1989	0.2007	0.1960	0.2280	0.2551	0.1148	0.0796	0.1034	0.0932	0.1515	0.2004	0.0260
	25ms	0.2134	0.2125	0.2075	0.2431	0.2733	0.1237	0.0952	0.1189	0.1070	0.1737	0.2283	0.0370
	50ms	0.2687	0.2614	0.2553	0.3018	0.3394	0.1514	0.1677	0.2089	0.1909	0.2882	0.3488	0.0709
	75ms	0.2920	0.2791	0.2729	0.3253	0.3668	0.1625	0.2039	0.2352	0.2160	0.3317	0.4006	0.0872
	100ms	0.3043	0.2887	0.2822	0.3381	0.3827	0.1692	0.2226	0.2470	0.2289	0.3548	0.4277	0.0975
	150ms	0.3197	0.3020	0.2955	0.3556	0.4031	0.1794	0.2491	0.2632	0.2480	0.3805	0.4609	0.1131
	200ms	0.3289	0.3103	0.3036	0.3664	0.4154	0.1857	0.2648	0.2727	0.2575	0.3979	0.4788	0.1232
	500ms	0.3519	0.3301	0.3225	0.3913	0.4445	0.2030	0.3028	0.2967	0.2798	0.4310	0.5152	0.1521
	750ms	0.3585	0.3351	0.3278	0.3982	0.4528	0.2097	0.3159	0.3028	0.2880	0.4387	0.5246	0.1644
	1sec	0.3617	0.3370	0.3293	0.4009	0.4559	0.2136	0.3219	0.3038	0.2894	0.4408	0.5266	0.1711
	5sec	0.3665	0.3410	0.3322	0.4059	0.4627	0.2278	0.3330	0.3080	0.2947	0.4341	0.5277	0.1966
	15sec	0.3645	0.3439	0.3337	0.4066	0.4641	0.2337	0.3349	0.3117	0.2995	0.4309	0.5246	0.2089
	30sec	0.3631	0.3467	0.3344	0.4082	0.4642	0.2378	0.3329	0.3185	0.3039	0.4314	0.5255	0.2159
	60sec	0.3636	0.3499	0.3367	0.4099	0.4647	0.2385	0.3396	0.3261	0.3125	0.4355	0.5249	0.2169
	300sec	0.3694	0.3640	0.3447	0.4160	0.4685	0.2398	0.3577	0.3539	0.3344	0.4459	0.5334	0.2187
	600sec	0.3741	0.3716	0.3496	0.4200	0.4719	0.2394	0.3674	0.3654	0.3439	0.4528	0.5378	0.2196
	900sec	0.3772	0.3756	0.3513	0.4204	0.4720	0.2380	0.3711	0.3714	0.3472	0.4557	0.5410	0.2187
	1200sec	0.3793	0.3779	0.3522	0.4215	0.4731	0.2398	0.3763	0.3753	0.3500	0.4576	0.5409	0.2203
	1500sec	0.3811	0.3805	0.3539	0.4224	0.4744	0.2393	0.3780	0.3786	0.3531	0.4586	0.5431	0.2194
	1800sec	0.3823	0.3819	0.3547	0.4226	0.4746	0.2392	0.3798	0.3811	0.3540	0.4595	0.5441	0.2207

Notes: This table reports both the mean and median of a series of liquidity spreads fifth-order autocorrelations for the full 10-year sample period, including the fifth-order autocorrelations of bid-ask spreads ($\Omega^{\tilde{\mathscr{F}}Q}$), quoted half-spread ($\Omega^{\tilde{\mathscr{F}}Q^{1/2}}$), quoted depth ($\Omega^{\tilde{\mathscr{F}}D}$), effective half-spread ($\Omega^{\tilde{\mathscr{F}}E}$), realized spreads ($\Omega^{\tilde{\mathscr{F}}R}$), and adverse selection ($\Omega^{\tilde{\mathscr{F}}AS}$). All the spreads autocorrelations are calculated by the volume weighted average (VWA) spreads and the spreads for each order book level. Both realized spreads autocorrelations and adverse selection autocorrelations are measured at 26 varying time intervals, including tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec. The fifth-order spreads autocorrelations for other tenor subsamples, such as 30 days to maturity and 1,800 days to maturity, are available in the online appendix.

Table 5.9: The Tenth-order Spreads Autocorrelation

		Mean						Median					
		VWA	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**	**VWA**	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
$\Omega^{\tilde{\mathscr{F}}Q}$		0.4625	0.1518	0.1671	0.2360	0.2605	0.2493	0.4874	0.1116	0.1226	0.1901	0.2146	0.1236
$\Omega^{\tilde{\mathscr{F}}Q^{1/2}}$		0.4625	0.1518	0.1671	0.2360	0.2605	0.2493	0.4874	0.1116	0.1226	0.1901	0.2146	0.1236
$\Omega^{\tilde{\mathscr{F}}D}$		0.4067	0.4084	0.6530	0.4960	0.7765	0.6110	0.3348	0.4060	0.7929	0.5082	0.8517	0.6816
$\Omega^{\tilde{\mathscr{F}}E}$		0.3140	0.2169	0.2064	0.2674	0.3108	0.1120	0.3056	0.1711	0.1508	0.2463	0.3110	0.0317
	Tick	0.3125	0.1870	0.1820	0.2455	0.2872	0.1255	0.1205	0.1053	0.2061	0.2694	0.0277	0.0000
	1ms	0.3139	0.2171	0.2065	0.2676	0.3111	0.1117	0.3055	0.1713	0.1509	0.2467	0.3116	0.0312
	2ms	0.3138	0.2172	0.2066	0.2678	0.3113	0.1120	0.3055	0.1715	0.1510	0.2471	0.3125	0.0311
	5ms	0.3135	0.2176	0.2071	0.2682	0.3121	0.1131	0.3049	0.1714	0.1511	0.2472	0.3136	0.0320
	10ms	0.3130	0.2183	0.2078	0.2691	0.3133	0.1140	0.3038	0.1726	0.1516	0.2495	0.3158	0.0334
	15ms	0.3125	0.2190	0.2086	0.2700	0.3146	0.1137	0.3031	0.1732	0.1529	0.2507	0.3179	0.0350
	20ms	0.3121	0.2197	0.2096	0.2710	0.3157	0.1133	0.3025	0.1735	0.1545	0.2519	0.3194	0.0357
	25ms	0.3118	0.2204	0.2104	0.2720	0.3166	0.1133	0.3020	0.1741	0.1553	0.2527	0.3210	0.0366
	50ms	0.3101	0.2247	0.2151	0.2774	0.3231	0.1145	0.2983	0.1780	0.1596	0.2600	0.3297	0.0436
	75ms	0.3084	0.2279	0.2184	0.2812	0.3276	0.1152	0.2961	0.1793	0.1619	0.2643	0.3360	0.0480
	100ms	0.3072	0.2297	0.2207	0.2839	0.3307	0.1164	0.2947	0.1805	0.1642	0.2675	0.3412	0.0512
	150ms	0.3056	0.2326	0.2238	0.2878	0.3358	0.1192	0.2926	0.1816	0.1645	0.2711	0.3486	0.0571
$\Omega^{\tilde{\mathscr{F}}R}$	**200ms**	0.3044	0.2348	0.2260	0.2907	0.3393	0.1220	0.2906	0.1820	0.1655	0.2734	0.3535	0.0615
	500ms	0.3016	0.2414	0.2334	0.3000	0.3500	0.1348	0.2863	0.1849	0.1697	0.2836	0.3716	0.0789
	750ms	0.3007	0.2442	0.2365	0.3036	0.3544	0.1413	0.2841	0.1866	0.1724	0.2875	0.3782	0.0871
	1sec	0.3005	0.2455	0.2380	0.3057	0.3568	0.1470	0.2830	0.1876	0.1732	0.2893	0.3808	0.0936
	5sec	0.3004	0.2533	0.2468	0.3165	0.3685	0.1828	0.2752	0.1925	0.1781	0.2988	0.3913	0.1372
	15sec	0.2971	0.2584	0.2526	0.3233	0.3763	0.2110	0.2664	0.1966	0.1849	0.3060	0.3998	0.1692
	30sec	0.2923	0.2605	0.2553	0.3262	0.3791	0.2297	0.2582	0.2006	0.1898	0.3078	0.4031	0.1935
	60sec	0.2886	0.2627	0.2575	0.3282	0.3815	0.2446	0.2530	0.2079	0.1968	0.3124	0.4043	0.2126
	300sec	0.2862	0.2720	0.2660	0.3353	0.3885	0.2695	0.2555	0.2297	0.2201	0.3262	0.4122	0.2371
	600sec	0.2881	0.2781	0.2716	0.3391	0.3915	0.2742	0.2591	0.2431	0.2332	0.3354	0.4182	0.2457
	900sec	0.2896	0.2813	0.2740	0.3409	0.3933	0.2750	0.2624	0.2494	0.2389	0.3386	0.4210	0.2479
	1200sec	0.2907	0.2836	0.2762	0.3426	0.3941	0.2802	0.2654	0.2532	0.2428	0.3422	0.4215	0.2512
	1500sec	0.2914	0.2856	0.2786	0.3431	0.3946	0.2764	0.2670	0.2564	0.2471	0.3430	0.4224	0.2493
	1800sec	0.2921	0.2871	0.2800	0.3442	0.3953	0.2780	0.2678	0.2587	0.2492	0.3441	0.4243	0.2501
	Tick	0.0980	-0.0007	-0.0004	0.0017	0.0020	0.0016	0.0013	0.0000	0.0000	0.0000	-0.0001	0.0000
	1ms	0.0658	0.0196	0.0193	0.0203	0.0211	0.0005	0.0190	-0.0003	-0.0003	-0.0002	-0.0002	-0.0001
	2ms	0.0590	0.0332	0.0327	0.0353	0.0378	0.0042	0.0104	-0.0003	-0.0003	-0.0002	-0.0002	-0.0001
	5ms	0.0696	0.0637	0.0626	0.0703	0.0775	0.0196	0.0077	-0.0002	-0.0002	-0.0001	-0.0001	-0.0001
	10ms	0.0935	0.0913	0.0898	0.1037	0.1161	0.0416	0.0100	-0.0001	-0.0001	0.0032	0.0203	-0.0001
	15ms	0.1115	0.1083	0.1063	0.1251	0.1410	0.0547	0.0159	0.0008	0.0000	0.0241	0.0468	-0.0001
	20ms	0.1291	0.1244	0.1222	0.1446	0.1638	0.0654	0.0224	0.0168	0.0131	0.0444	0.0756	0.0000
	25ms	0.1417	0.1342	0.1318	0.1573	0.1789	0.0716	0.0288	0.0282	0.0233	0.0606	0.0961	0.0000
	50ms	0.1923	0.1746	0.1720	0.2074	0.2363	0.0936	0.0659	0.0809	0.0716	0.1432	0.1946	0.0136
	75ms	0.2169	0.1917	0.1890	0.2302	0.2630	0.1016	0.0917	0.1014	0.0906	0.1805	0.2369	0.0224
	100ms	0.2299	0.2005	0.1979	0.2421	0.2779	0.1065	0.1060	0.1122	0.1019	0.1980	0.2623	0.0275
	150ms	0.2457	0.2117	0.2094	0.2578	0.2966	0.1135	0.1285	0.1265	0.1150	0.2215	0.2944	0.0356
$\Omega^{\tilde{\mathscr{F}}AS}$	**200ms**	0.2551	0.2183	0.2161	0.2669	0.3072	0.1176	0.1409	0.1347	0.1248	0.2350	0.3114	0.0410
	500ms	0.2780	0.2355	0.2334	0.2905	0.3354	0.1302	0.1749	0.1573	0.1471	0.2673	0.3495	0.0612
	750ms	0.2841	0.2409	0.2389	0.2980	0.3443	0.1354	0.1845	0.1643	0.1534	0.2774	0.3626	0.0692
	1sec	0.2873	0.2428	0.2408	0.3011	0.3479	0.1379	0.1895	0.1650	0.1554	0.2795	0.3668	0.0749
	5sec	0.2895	0.2468	0.2445	0.3084	0.3578	0.1500	0.2005	0.1728	0.1636	0.2788	0.3719	0.0991
	15sec	0.2847	0.2496	0.2468	0.3110	0.3613	0.1563	0.2046	0.1813	0.1724	0.2817	0.3735	0.1131
	30sec	0.2791	0.2513	0.2473	0.3124	0.3627	0.1604	0.2052	0.1874	0.1766	0.2833	0.3744	0.1200
	60sec	0.2758	0.2538	0.2490	0.3148	0.3647	0.1622	0.2103	0.1977	0.1871	0.2901	0.3765	0.1237
	300sec	0.2766	0.2674	0.2585	0.3241	0.3729	0.1672	0.2333	0.2282	0.2116	0.3102	0.3922	0.1302
	600sec	0.2805	0.2742	0.2633	0.3287	0.3764	0.1675	0.2441	0.2406	0.2236	0.3193	0.4022	0.1308
	900sec	0.2825	0.2779	0.2653	0.3293	0.3772	0.1667	0.2485	0.2471	0.2286	0.3219	0.4023	0.1306
	1200sec	0.2843	0.2804	0.2667	0.3308	0.3790	0.1684	0.2536	0.2519	0.2308	0.3237	0.4035	0.1326
	1500sec	0.2857	0.2826	0.2682	0.3313	0.3800	0.1675	0.2571	0.2553	0.2337	0.3268	0.4059	0.1313
	1800sec	0.2866	0.2841	0.2690	0.3319	0.3802	0.1674	0.2588	0.2572	0.2358	0.3268	0.4067	0.1324

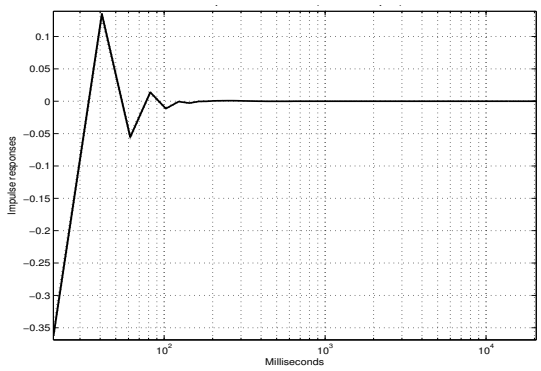
Notes: This table reports both the mean and median of a series of liquidity spreads tenth-order autocorrelations for the full 10-year sample period, including the tenth-order autocorrelations of bid-ask spreads ($\Omega^{\tilde{\mathscr{F}}Q}$), quoted half-spread ($\Omega^{\tilde{\mathscr{F}}Q^{1/2}}$), quoted depth ($\Omega^{\tilde{\mathscr{F}}D}$), effective half-spread ($\Omega^{\tilde{\mathscr{F}}E}$), realized spreads ($\Omega^{\tilde{\mathscr{F}}R}$), and adverse selection ($\Omega^{\tilde{\mathscr{F}}AS}$). All the spreads autocorrelations are calculated by the volume weighted average (VWA) spreads and the spreads for each order book level. Both realized spreads autocorrelations and adverse selection autocorrelations are measured at 26 varying time intervals, including tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec. The tenth-order spreads autocorrelations for other tenor subsamples, such as 30 days to maturity and 1,800 days to maturity, are available in the online appendix.

red line, three-year tenor with the blue line, and five-year tenor with the black line in the Figure 5.11. Moreover, I implement the 95% confident intervals for the impulse response via i.i.d. bootstrap with 99 resamples. The dotted lines with corresponding colors are the 95% lower and upper confidence bounds for the impulse response functions.

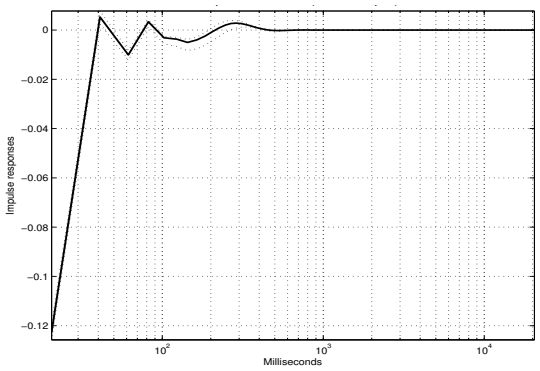
Here I will not report all of the impulse response plots in this chapter, but select the trade returns impulse responses. To a certain degree, the adjustment of trade return directly reflects the price discovery and the market efficiency. However, the impulse responses of a signed trade direction indicator, signed trade volume, signed square root of trade volume, signed bid-side and ask-side market concentration ratio are available in the online appendix.

Figure 5.10 plots the impulse response of trade return to all six factors for the full sample period. Subplot (a) provides a benchmark adjustment because it is the impulse response function in trade return to shocks to itself. The results indicate that trade return has an effect on itself before 200 milliseconds in the full sample. As the benchmark of the response of the trade return, the effect initially is negative below 35 ms. This means around 35 ms if a one standard deviation of shock is given to the trade return, it will decrease 0.37 units. Between 35 ms and 55 ms, the response in trade return becomes a positive shift to a shock to itself.

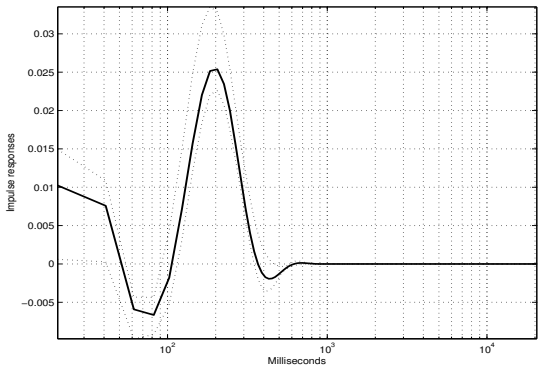
When a one standard deviation shock is given to the trade return, the largest positive impact occurs at 40 ms as the benchmark of the response of trade return. Additionally, 40 milliseconds appears also to be a critical point for the trade return adjustment to the shock in signed trade direction indicators for the full sample dataset (see Figure 5.10 Subplot (b)). One unit standard deviation shock from the trade direction indicator will cause 0.12 standard deviation units negatively impacts on trade return around 20 ms, which is smaller than the reaction of



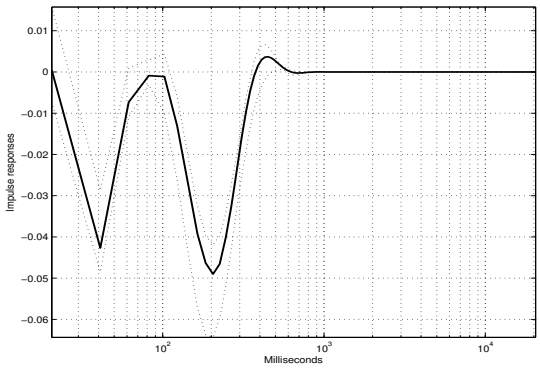
(a) Response of Trade Return to Trade Return



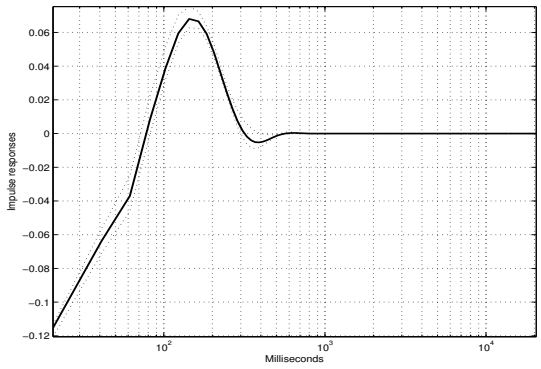
(b) Response of Trade Return to Trade Direction



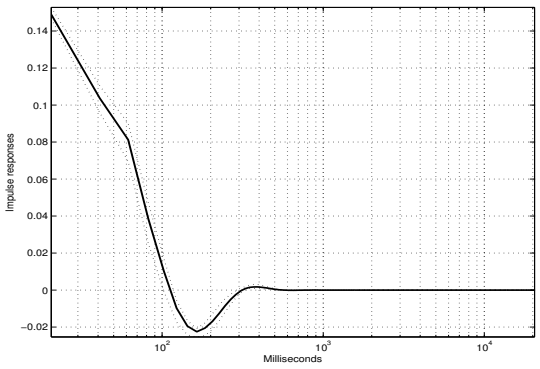
(c) Response of Trade Return to Trade Volume



(d) Response of Trade Return to Square Root Trade Volume



(e) Response of Trade Return to Asks Market Concentration



(f) Response of Trade Return to Bids Market Concentration

Figure 5.10: Response of Trade Return (Full Sample)

Notes: This figure portrays the impulse response of trade returns to the shocks in the trade return, signed trade direction, signed trade volume, signed square root of trade volume, and signed asks and bids market concentration ratio for the full sample dataset across the 10-year life cycle of 40 Eurodollar contracts. The impulse response functions are conducted with the first-order vector autoregression with 1,000 steps on the tick-by-tick basis data, and there are approximately 178 million observations in total. The black line represents the impulse response of trade return to the shocks in each variables, and the dotted line are the 95% lower and upper confidence bounds of the impulse response functions via i.i.d bootstrap with 99 resampling bootstrap procedure.

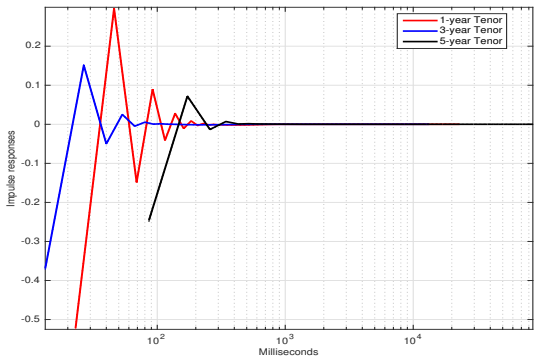
trade return to itself. The effect of trade return dramatically rises to the top at 40 milliseconds with positive impacts from the trader direction indicator on trade return. After a half second, signed trade direction indicators do not impact on trade return.

In Figure 5.10 Subplots (c) and (d), the signed trade volume and the signed square root of trade volume affect the trade return in a very short timeframe. The trade return increases with a shock to the signed trade volume at the beginning, with one order of magnitude smaller than the response of the trade return to its own shock. Specifically, the signed trade volume starts affecting trade return around 20 ms with 0.01 units positive influence on return (see Figure 5.10 Subplot (c)). From 50 ms to 110 ms, a one standard deviation shock to trade volume results in instant negative shifts in trade return. The lowest point is around 80ms with negatively 0.007 units impacts on trade return. The effects of the shock from trade volume positively increase after 110 milliseconds, and reach a peak around 200 ms with a 0.025 units increase of the trade return. Once it passes 700 milliseconds, trades volume has no impacts on the trade return. In Figure 5.10 Subplot (d), the shock to the square root of trade volume starts to have adverse effects on trade return from 20 ms to 370 ms. One standard deviation shock to the square root of trade volume results in -0.043 standard deviation units shifts in the trade return at 40 milliseconds. When it comes to 200 milliseconds, the negative effects on trade return reach to -0.049. After 700 milliseconds, the square root of trade volume is not affecting the trade return anymore.

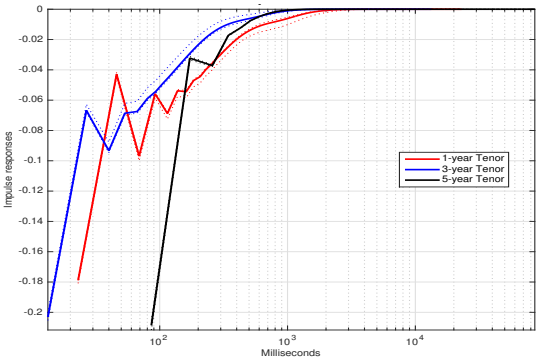
Both signed ask-side and bid-side market concentration ratios also have influences on the trade return (see Figure 5.10 Subplot (e) and (f)). In Figure 5.10 Subplot (e), the effect of the shock to signed ask-side market concentration starts around 20 milliseconds with negatively 0.115 units impacts in the trade return.

Between 80 ms and 300 ms, a one standard deviation of shock to the ask-side market concentration results in the positive increase of the trade return, with the highest response of trade return about 0.07 at 150 milliseconds. Once it passes 600 milliseconds, the ask-side market concentration shows no instant impacts on the trade return. For the response of trade return from the bid-side market concentration, it begins with a 0.15 increase in the trade return at 20 milliseconds (see Figure 5.10 Subplot (f)). The trade return decreases with a shock to the signed bid-side market concentration ratio from 120 milliseconds, and reaches the bottom with -0.021 standard deviation units shift at 170 milliseconds. After 500 milliseconds, the signed bid-side market concentration stops affecting the trade return.

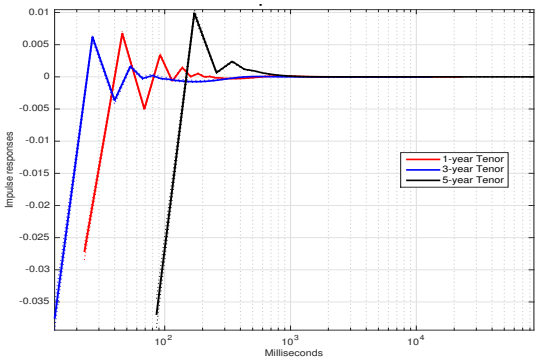
To interpret the influence of the days to maturity date on the response of trade returns, Figure 5.11 presents the impulse response functions of trade return to six factors with one-year tenor, three-year tenor, and five-year tenor Eurodollar contracts. In Figure 5.11 Subplot (a), the three-year tenor contract with blue line is the earliest one reacting to the shift from trade return, starting from -0.36 at ten milliseconds, crossing the zero line at 21 milliseconds and reaching the top with 0.15 at 27 milliseconds. The one-year to maturity subsample with the red line, however, shows the highest response of trade return with the -0.52 around 23 milliseconds and the 0.29 around 46 milliseconds. This suggests that if the trade return will increase one unit standard deviation, it will decrease 0.52 standard deviation units after 23 ms and will increase 0.29 units after 46 milliseconds. The black line indicates that the Eurodollar contracts within the five-year to maturity has slower speed and lower response of trade return to itself, compared with the response of trade return with one-year and three-year tenors. Hence, Figure 5.11 Subplot (a) implies that trade return within one year to maturity period has the largest impact on itself, but the fastest trade return adjustment is within the



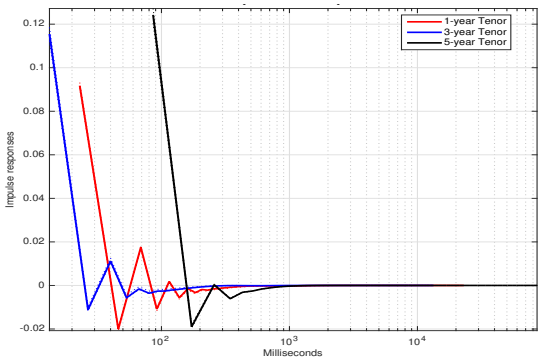
(a) Response of Trade Return to Trade Return



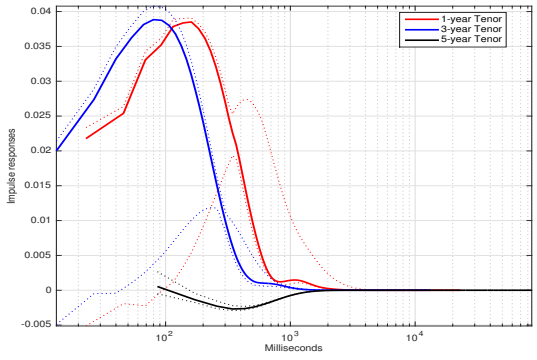
(b) Response of Trade Return to Trade Direction



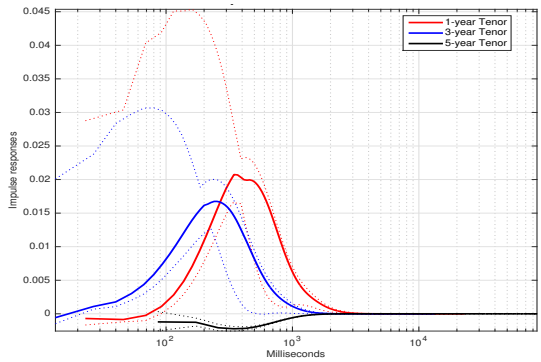
(c) Response of Trade Return to Trade Volume



(d) Response of Trade Return to Square Root Trade Volume



(e) Response of Trade Return to Asks Market Concentration



(f) Response of Trade Return to Bids Market Concentration

Figure 5.11: Response of Trade Return (First-year, Third-year and Fifth-year Tenor)

Notes: This figure portrays the impulse response of trade returns to the shocks in the trade return, signed trade direction, signed trade volume, signed square root of trade volume, and signed asks and bids market concentration for the three subsamples with one-year tenor, three-year tenor, and five-year tenor of 40 Eurodollar contracts. The impulse response functions are conducted with the first-order vector autoregression with 1,000 steps on the tick-by-tick basis data. The red line represents the response of trade return of one-year tenor subsample with around 43.5 million observations, the blue line is the three- year tenor impulse response of trade return with 42.5 million observations, and the black line is the five-year tenor impulse response of trade return with 10.4 million observations. The dotted lines with corresponding colours are the 95% lower and upper confidence bounds of the impulse response functions via i.i.d bootstrap with 99 resampling bootstrap procedure.

three-year to maturity time-window.

Subplot (b) of Figure 5.11 illustrates the impulse response of trade return to a unit shock in the trade direction indicator with the one-year, three- year, and five-year to maturity. Trade return appears to respond negatively to trade direction indicator for all three different datasets. The shock in the trade direction indicator starts affecting the trade return from 23 milliseconds with -0.18 standard deviation units shift for the one-year to maturity group. The negative response of trade return decreases and drops back to zero at two seconds. The speed of trade return adjustment in the three-year tenor is faster than others, with the shortest reaction time ten milliseconds with -0.2. This suggests a one standard deviation of trade direction will cause the trade return to decrease 0.2 units with ten milliseconds adjustment timeframe. For the five-year to maturity dataset, the response of trade return to trade direction indicators begins at 85 milliseconds. The trade direction stops influencing trade return for both the three-year and five-year tenor datasets, once the adjustment timeframe is longer than one second.

Figure 5.11 Subplots (c) and (d) present the responses of trade return to the signed trade volume and the signed square root of trade volume in three different subsamples. The largest effects from the trade volume to trade return are the -0.038 at ten milliseconds for the three-year tenor contracts, -0.027 at 23 milliseconds for one-year tenor contracts and -0.037 at 85 milliseconds for the five-year tenor contracts (see Figure 5.11 Subplot (c)). Three lines, represented as three different tenor contracts, converge together to zero around 900 milliseconds, which implies that the trade volume has no impacts on trade return after 900 milliseconds. In Figure 5.11 Subplot (d), the largest response of trade return in facing the shock from the square root of trade volume is in the five-year tenor

contract dataset with a 0.125 units increase on trade return around 85 milliseconds. For the one-year tenor contracts, a one standard deviation shock to the square root of trade volume results in a 0.091 units increasing shift of trade return at 23 milliseconds. Furthermore, the response of trade return in the third-year tenor group has a 0.115 positively shift around ten milliseconds from the shock to the square root of trade volume. Similarly with Subplot (c), the response of trade return to the square root of trade volume seems to stop after 900 milliseconds.

Additionally, ask and bid market concentrations also have the instant impacts on the trade returns in three subsamples (see Figure 5.11 Subplots (e) and (f)). At 80 milliseconds, one standard deviation of shock in the ask-side market concentration will cause the one-year tenor trade return increasing by 0.0385 standard deviation units. For the three-year to maturity contracts, if the ask-side market concentration rises one unit, trade return will increase by 0.038 units at 160 milliseconds as the largest response to trade return. The ask-side market concentration in the five-year to maturity contracts has positive effects from 85 ms to 100 ms, and drops to the bottom with -0.0025 around 400 milliseconds. The impacts from the bid-side market concentration on the trade return are lower than the impacts from the ask-side (see Figure 5.11 Subplots (f)). The highest response of trade return is in the one-year to maturity contracts with 0.021 standard deviation units shift at 350 milliseconds. For the three-year tenor contracts, one standard deviation shock in the bid-side market concentration leads to the trade return increasing by 0.017 standard deviation units around 250 milliseconds. However, the bid-side market concentration has negative impacts on the trade return for five-year tenor contracts between 85 milliseconds and 1.75 seconds. Compared with the response in one-year and three-year tenor contracts, the response of trade return to both-side market concentrations in the five-year tenor have much tighter 95% confidence intervals with more significant impulse

response results.

It is apparent from the results that all the six factors have statistically significant effects on the trade return; however the direction and dynamics vary. For the whole cross section of 10-year to overnight Eurodollar tenors, the largest trade return shift is triggered by the shock in itself around ten milliseconds, with 0.35 standard deviation units decreasing of trade return. Moreover, the fastest adjustment of trade return is also from the shock of itself. The whole adjustments are finished within 200 milliseconds timeframe. The slowest trade returns adjustments are within 700 milliseconds, which are the responses of trade return to the shock in both the trade volume and the square root of trade volume. This suggests the price discovery process is very fast and instant, and is likely to finish within one second. Furthermore, there is strong evidence of the critical turning points in the response of trade returns to the shocks in diverse market indicators. To capture a wide variety of the Eurodollar future market properties, I also measure the impulse responses functions of all six factors for three different tenor Eurodollar contracts. The results are consistent and complement the 10-year contracts life-cycle. It is worth noting that the shocks in the trade direction indicators result in negative direction adjustment in the trade return for all three different tenor contracts. Except for the trade direction indicators, other five factors show the existence of turning points for the trade return adjustment. Moreover, the one-year tenor contract has the greatest response, absolute magnitude, in trade return, and the three-year tenor contracts are the earliest one affecting the trade return.

5.4.3 Regression Interpretation and Robustness Checks

To provide a complementary understanding of the price discovery, this chapter considers the maturity effects as an important factor for the efficiency of price adjustments in the future market. Some prior literature has found it is a challenge to find the maturity effects on the future market, such as Bessembinder et al. [1996] and Kalev and Duong [2008]. However, my empirical evidence supports Samuelson's maturity effect theory [Samuelson, 1965] and even suggests that maturity effects do impact the price efficiency in Eurodollar future market. To investigate the effectiveness of price discovery can be affected by the time to maturity, I utilize a multivariate linear regression to verify the maturity effects on the price discovery and market quality.

The abundance of data allows for a large number of seemingly unrelated regression to be estimated. The regression includes 319 dependent variables, mainly from price discovery factors and market quality factors. The independent variables cover the days to maturity ($\tilde{T}$), the inverse of transaction prices (p^{-1}), and the daily volatility ($\mathcal{V}$). In all cases, robust standard errors are utilized. All variables are daily volume weighted average with 32,712 days across all 40 contracts. Besides, I also provide the estimations on three different subsamples based on the days before settlement dates, namely one year to maturity group ($\tilde{T}$ smaller than one year), three years to maturity group ($\tilde{T}$ smaller than three years), and five years to maturity group ($\tilde{T}$ smaller than five years). I report the coefficient estimates of the days to maturity in Table 5.10 to Table 5.12, other estimations and statistics are available in the online appendix.

Table 5.10 reports the coefficients of the days to maturity on the market activities indicators (the number of trades and quotes, the quote-to-trade ratio) and the

Table 5.10: Market Activities and Price Efficiency Regression Results

		Full Sample	1-Year Tenor	3-Year Tenor	5-Year Tenor
Panel A: Market Activities					
Number of quotes (n^q)		0.3325***	0.5795***	0.5335***	0.4753***
Number of trades (n)		-0.2010***	0.2102***	0.1374***	0.0612***
Quote-to-trade ratio ($\mathcal{Q}$)		0.5335***	0.3693***	0.3961***	0.4141***
Panel B: Quotes Price Autocorrelations (Ω^P)					
First-order	Ask	-0.0164***	-0.0006	-0.0029***	0
	Bid	-0.0259***	-0.0015**	-0.0032***	-0.0045***
	Mid	-0.0222***	-0.001	-0.0024***	-0.0019**
Fifth-order	Ask	-0.0477***	-0.0029*	-0.0091***	-0.0035**
	Bid	-0.0593***	-0.0043**	-0.0095***	-0.0104***
	Mid	-0.0554***	-0.0037**	-0.0091***	-0.0069***
Tenth-order	Ask	-0.0652***	-0.0038*	-0.0123***	-0.0058***
	Bid	-0.0765***	-0.0051**	-0.0132***	-0.0137***
	Mid	-0.0731***	-0.0045**	-0.0130***	-0.0101***
Panel C: Pricing Error ($\mathcal{E}$)					
Pricing error ($\mathcal{E}$)	Lag 1	0.0521***	0.1631***	0.0202	-0.0755***
	Lag 2	0.1530***	0.2403***	0.0594	-0.0653**
	Lag 3	0.2220***	0.2127***	0.0880**	-0.0202
	Lag 4	0.1866***	0.1953***	0.0665*	-0.0447
	Lag 5	0.1404***	0.0219	-0.0402	-0.1188***

* p<0.1; ** p¡0.05; *** p<0.01

Notes: This table describes the coefficients of the days to maturity ($\tilde{T}$) on a set of market activities indicators and price discovery factors for the 10-year full sample period, one year tenor group ($\tilde{T}$ smaller than one year), three years tenor group ($\tilde{T}$ smaller than three years), and five years tenor group ($\tilde{T}$ smaller than five years). I utilize the following multivariate linear regression model to verify the maturity effects on the price discovery and market quality, $\tilde{\boldsymbol{Y}} = \boldsymbol{c}+\tilde{\boldsymbol{T}}+\boldsymbol{p^{-1}}+\boldsymbol{\mathcal{V}}$, where explanatory variables include the days to maturity ($\tilde{T}$), the inverse of transaction prices (p^{-1}), and the daily volatility ($\mathcal{V}$) is the highest daily transaction price subtracting the lowest daily transaction price. Dependent variables in the Panel A includes the number of quotes and trades (n^q and n) and the quotes to trades ratio $\mathcal{Q}$. Dependent variables in the Panel B are the first-order, fifth-order, and tenth-order autocorrelations (Ω^P) of the volume weighted average bid, ask, and mid-quote price at tick-by-tick level. Dependent variables in the Panel C are the pricing errors ($\mathcal{E}$) measured by the lag one to lag five vector autoregression models. It is worth noting that trades number, quotes number, pricing errors, the inverse of price, the days to maturity and the quotes to trades ratio take the logarithm value. All variables are the daily volume-weighted average results across five level order book data, with 32,712 observations in the 10-year full sample, 4,683 daily observations in the one year group, 15,009 daily observations in the three years group, and 25,481 daily observations in the five years group across 40 Eurodollar contracts. The asterisks $***$, $**$, and $*$ denote 1%, 5%, and 10% statistical significance of the coefficients of days to maturity. This table only reports the coefficient estimates of the days to maturity, other estimations and statistics are available in the online appendix.

price discovery factors (pricing errors and the autocorrelations of the volume weighted bids, asks, and mid quotes). Panel A suggests the intensity of quote submission increases since the contract approaches the maturity date. From year five to year one, the coefficient estimates of quote submission rise from 0.4753 to 0.5795, while the full sample coefficient is only 0.3325, suggesting that, after year five, the Eurodollar market is less active. It is worth noting that the coefficient of the trade number in the full sample group is significantly negative related to the time to maturity; however, it is positively decreasing from year one group to year five group. Combined with the decreasing tendency of the quote-to-trade ratio from 0.4141 at year five group to 0.3693 in year one, I conclude that when the contract moves to its settlement date, the quotes heavily grow, but, instead of increasing the number of transactions, traders are more likely to increase their trading volumes dramatically. Table 5.10 Panel B illustrates an inverse relation between the distance to maturity and the volume weighted quotes autocorrelations. The closer to maturity, the larger number of quotes submitted to the limit order book. Hence, both bid-side and ask-side prices are highly autocorrelated near maturity.

To assess the maturity effects on the price discovery process in the Eurodollar market, I follow the Hasbrouck [1993] vector autoregression method to calculate the daily pricing error for each contract, and measure the mid-price return autocorrelations using Hendershott and Jones [2005] (see Table 5.10 Panel C and Table 5.11). The results suggest that the distance to maturity date has positively significant effects on the pricing error. The coefficient of maturity date in lag one pricing error is 0.0521, and the highest coefficient is in lag three groups with 0.2220. However, there are also negative coefficients at five years to maturity subsample. The estimation indicates that the pricing error will increase when the contract approaches its maturity. In the lag-one pricing error group, the

coefficient estimates rise from -0.0755 in five years to maturity group to 0.1631 in one year to maturity group. Similar changes can also be found in other lags pricing error estimation. Because the pricing error reflects the deviation of the trade prices from the efficient prices, my results demonstrate that since the future contracts approach maturity, the transaction prices are less sensitive absorbing available information in the market and are farther away from the actual value of the contract. Because of the extremely active trading near the maturity, the amount of information and inside trading also increases, which could add the difficulty of incorporating information and efficiently adjust the prices.

In Table 5.11, I report the coefficient of maturity date on the first-order, fifth-order, and tenth-order volume weighted mid-price return autocorrelations at 26 time intervals from tick level to half hour, respectively. In an efficient market, the level of return autocorrelation should be tending to zero, so the reason Table 5.11 is interesting it tells you that what timescale you are looking at an efficient market, and what timescale you are looking at a non-efficient market.

The coefficient of the days to maturity with the first-order, fifth-order and tenth-order return autocorrelations all peak around 1ms and decrease to 1800 seconds. Specifically, the coefficient of the days to maturity with the first-order return autocorrelation is 0.0104 at tick level and reaches the peak with 0.0966 at 1 ms. This suggests that the distance to maturity has the largest impacts when the first-order mid-quote return autocorrelation is measured at 1 ms time intervals. Once the timescale is longer than 1 ms, the coefficient starts to decline from 0.0712 at 2ms to -0.0296 at 30 minutes. When the timescale is longer than 750 ms, the coefficients of the days to maturity are below zero, which implies that the days to maturity has negative effects on the price efficiency.

For each subsample, the finding indicates that the first-order return autocorrela-

Table 5.11: Mid-quote Return Autocorrelation Regression Results

	First-order				Fifth-order				Tenth-order			
	Full Sample	1-Year Tenor	3-Year Tenor	5-Year Tenor	Full Sample	1-Year Tenor	3-Year Tenor	5-Year Tenor	Full Sample	1-Year Tenor	3-Year Tenor	5-Year Tenor
Tick	0.0104***	0.0100***	0.0188***	0.0177***	0.0024***	0.0023***	0.0042***	0.0047***	0.0003	0.0007	0.0014***	0.0047***
1ms	0.0966***	0.0639***	0.1064***	0.0995***	0.0272***	0.0168***	0.0326***	0.0296***	0.0117***	0.0069***	0.0155***	0.0296***
2ms	0.0712***	0.0737***	0.0837***	0.0780***	0.0166***	0.0240***	0.0285***	0.0218***	0.0067***	0.0135***	0.0145***	0.0218***
5ms	0.0386***	0.0583***	0.0598***	0.0543***	0.0080***	0.0357***	0.0326***	0.0197***	0.0021*	0.0201***	0.0181***	0.0197***
10ms	0.0466***	0.0462***	0.0679***	0.0667***	0.0089***	0.0385***	0.0411***	0.0259***	-0.0002	0.0257***	0.0238***	0.0259***
15ms	0.0500***	0.0517***	0.0758***	0.0739***	0.0134***	0.0436***	0.0502***	0.0338***	0.001	0.0323***	0.0304***	0.0338***
20ms	0.0506***	0.0525***	0.0781***	0.0761***	0.0164***	0.0480***	0.0571***	0.0399***	0.0018	0.0361***	0.0352***	0.0399***
25ms	0.0537***	0.0530***	0.0817***	0.0807***	0.0201***	0.0515***	0.0637***	0.0457***	0.0031**	0.0385***	0.0397***	0.0457***
50ms	0.0423***	0.0515***	0.0777***	0.0766***	0.0228***	0.0694***	0.0830***	0.0586***	0.0039**	0.0595***	0.0588***	0.0586***
75ms	0.0323***	0.0478***	0.0710***	0.0696***	0.0219***	0.0736***	0.0896***	0.0628***	0.0036**	0.0674***	0.0675***	0.0628***
100ms	0.0255***	0.0474***	0.0659***	0.0637***	0.0208***	0.0771***	0.0926***	0.0641***	0.0033*	0.0731***	0.0728***	0.0641***
150ms	0.0179***	0.0478***	0.0610***	0.0581***	0.0188***	0.0803***	0.0946***	0.0653***	0.003	0.0799***	0.0790***	0.0653***
200ms	0.0131***	0.0479***	0.0580***	0.0546***	0.0179***	0.0816***	0.0955***	0.0660***	0.0035*	0.0835***	0.0832***	0.0660***
500ms	0.0033	0.0459***	0.0497***	0.0459***	0.0145***	0.0829***	0.0952***	0.0671***	0.0056***	0.0886***	0.0922***	0.0671***
750ms	-0.0005	0.0428***	0.0457***	0.0422***	0.0119***	0.0819***	0.0934***	0.0661***	0.0050**	0.0881***	0.0936***	0.0661***
1sec	-0.0039*	0.0405***	0.0427***	0.0391***	0.0094***	0.0786***	0.0907***	0.0644***	0.0044**	0.0856***	0.0933***	0.0644***
5sec	-0.0172***	0.0322***	0.0301***	0.0263***	-0.0060***	0.0691***	0.0747***	0.0516***	-0.0034*	0.0810***	0.0856***	0.0516***
15sec	-0.0240***	0.0238***	0.0214***	0.0171***	-0.0203***	0.0611***	0.0590***	0.0374***	-0.0164***	0.0739***	0.0720***	0.0374***
30sec	-0.0286***	0.0212***	0.0164***	0.0123***	-0.0324***	0.0574***	0.0479***	0.0267***	-0.0273***	0.0724***	0.0615***	0.0267***
60sec	-0.0321***	0.0188***	0.0124***	0.0082***	-0.0432***	0.0515***	0.0367***	0.0169***	-0.0395***	0.0686***	0.0501***	0.0169***
300sec	-0.0349***	0.0088***	0.0045***	0.0021*	-0.0615***	0.0334***	0.0174***	0.0046**	-0.0644***	0.0511***	0.0270***	0.0046**
600sec	-0.0336***	0.0063***	0.0020**	0.001	-0.0657***	0.0244***	0.0094***	0.0019	-0.0710***	0.0407***	0.0176***	0.0019
900sec	-0.0321***	0.0047***	0.001	0.0004	-0.0663***	0.0197***	0.0065***	0.001	-0.0735***	0.0330***	0.0126***	0.001
1200sec	-0.0311***	0.0036***	0.0006	-0.0001	-0.0667***	0.0158***	0.0040**	-0.0004	-0.0748***	0.0271***	0.0086***	-0.0004
1500sec	-0.0304***	0.0031***	0.0004	-0.0003	-0.0663***	0.0139***	0.0027*	-0.0006	-0.0755***	0.0237***	0.0064***	-0.0006
1800sec	-0.0296***	0.0029***	0.0001	-0.0004	-0.0659***	0.0128***	0.0017	-0.0008	-0.0758***	0.0215***	0.0045**	-0.0008

* p<0.1; ** p¡0.05; *** p<0.01

Notes: This table describes the coefficients of the days to maturity ($\tilde{T}$) on a set of the volume weighted average mid-quote return for the 10- year full sample period, one year to maturity group ($\tilde{T}$ smaller than one year), three years to maturity group ($\tilde{T}$ smaller than three years), and five years to maturity group ($\tilde{T}$ smaller than five years). I utilize the following multivariate linear regression model to verify the maturity effects on the price discovery and market quality, $\tilde{\boldsymbol{Y}} = \boldsymbol{c} + \tilde{\boldsymbol{T}} + \boldsymbol{p^{-1}} + \boldsymbol{\mathcal{V}}$, where explanatory variables include the days to maturity ($\tilde{T}$), the inverse of transaction prices (p^{-1}), and the daily volatility ($\mathcal{V}$) is the highest daily transaction price subtracting the lowest daily transaction price. Dependent variables covers the first-order, fifth-order, and tenth-order mid-quote return autocorrelations at 26 varying time intervals, including tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec. It is worth noting that the inverse of price and the days to maturity take the logarithm value. All variables are the daily volume-weighted average results across five level order book data, with 32,712 observations in the 10-year full sample, 4,683 daily observations in the one year group, 15,009 daily observations in the three years group, and 25,481 daily observations in the five years group across 40 Eurodollar contracts. The asterisks $***$, $**$, and $*$ denote 1%, 5%, and 10% statistical significance of the coefficients of days to maturity. This table only reports the coefficient estimates of the days to maturity, other estimations and statistics are available in the online appendix.

tion at tick level increases from 0.01 at one year to maturity sample to 0.0188 at three years to maturity sample, and then has a slight drop to 0.0177 at five years to maturity subsample. Both fifth-order and tenth-order return autocorrelation also show the similar trends. In other words, the mid-quote return autocorrelations of three-year tenor contracts are more sensible in terms of the changes of the days to maturity. Table 5.11 also provides the estimations of the impacts of maturity date on the fifth-order and tenth-order mid-price return autocorrelation. Consistent with the first-order return autocorrelation, the time interval of one second is the turning point. Below one second, the days to maturity has a positive effect on return autocorrelations; longer than one second, the relation inverse to negative. Overall, the mid-quote return autocorrelation has its periodicity and is significantly related to the days to maturity, which also demonstrates the maturity effects do impact on the price efficiency in the Eurodollar future market.

Since trading activity increases dramatically near the maturity date, I expect that the time to maturity (tenor) would have an effect on the trading costs via the market liquidity spreads (see Hendershott et al., 2011). Table 5.12 displays the estimation results of liquidity spreads. The spreads are measured as the daily volume weighted average in basis point, except that the quoted depth is taking the logarithm value. The days to maturity has negative effects on the bid-ask spread for the full sample. However, the coefficient of maturity days on bid-ask spread is 15.6391 for the one year subsample. This means that, when the days are closer to maturity, the bid-ask spread is much tighter suggesting better liquidity in the market for the one year subsample. The coefficient estimations are all negative for the other two subsamples. The coefficient of days to maturity on quoted depth, -0.5663, indicates that the quoted depth becomes deeper as the contract approaches the maturity date, which enhances the ability to absorb

Table 5.12: Market Liquidity Spreads Regression Results

		Full Sample	1-Year Tenor	3-Year Tenor	5-Year Tenor		Full Sample	1-Year Tenor	3-Year Tenor	5-Year Tenor
$\tilde{\mathscr{J}}^{\boldsymbol{Q}}$		-211.932*	15.6391***	-10.9107**	-43.1972***	$\tilde{\mathscr{J}}^{\boldsymbol{D}}$	-0.5663*	0.1662**	0.1741***	-0.1366***
$\tilde{\mathscr{J}}^{\boldsymbol{Q^{1/2}}}$		-105.9663*	7.8196***	-5.4553**	-21.5986***	$\tilde{\mathscr{J}}^{\boldsymbol{E}}$	-0.9794*	0.0128***	-0.0387***	-0.1171***
	Tick	1.5573***	0.2308***	-0.3777***	1.7691***		0.0500	-0.0121	0.0221***	0.0593
	1ms	-0.9794*	0.0127***	-0.0387***	-0.1171***		0.0000	0.0000***	0.0000***	0.0000***
	2ms	-0.9794*	0.0127***	-0.0387***	-0.1172***		0.0001	0.0000***	0.0001***	0.0000***
	5ms	-0.9792*	0.0127***	-0.0389***	-0.1173***		-0.0002	0.0001***	0.0002***	0.0002***
	10ms	-0.9786*	0.0126***	-0.0391***	-0.1176***		-0.0007	0.0002***	0.0004***	0.0005***
	15ms	-0.9741*	0.0125***	-0.0392***	-0.1177***		-0.0053	0.0003***	0.0006***	0.0006***
	20ms	-0.9735*	0.0124***	-0.0394***	-0.1178***		-0.0058	0.0004***	0.0008***	0.0007***
	25ms	-0.9685*	0.0123***	-0.0396***	-0.1179***		-0.0109	0.0004***	0.0009***	0.0008***
	50ms	-0.9616*	0.0119***	-0.0403***	-0.1175***		-0.0178	0.0008***	0.0016***	0.0004
	75ms	-0.9622*	0.0115***	-0.0409***	-0.1174***		-0.0172	0.0012***	0.0022***	0.0003
	100ms	-0.9572*	0.0112***	-0.0413***	-0.1171***		-0.0221	0.0016***	0.0026***	0.0000
	150ms	-0.9581*	0.0105***	-0.0419***	-0.1185***		-0.0213	0.0022***	0.0033***	0.0014
$\tilde{\mathscr{J}}^{\boldsymbol{R}}$	**200ms**	-0.9599*	0.0100***	-0.0424***	-0.1195***	$\tilde{\mathscr{J}}^{\boldsymbol{AS}}$	-0.0195	0.0027***	0.0037***	0.0024***
	500ms	-0.9958*	0.0082***	-0.0438***	-0.1199***		0.0164	0.0046***	0.0051***	0.0028***
	750ms	-1.0017*	0.0073***	-0.0442***	-0.1197***		0.0223	0.0055***	0.0056***	0.0026***
	1sec	-0.9931*	0.0069***	-0.0445***	-0.1193***		0.0137	0.0059***	0.0059***	0.0022**
	5sec	-0.9917*	0.0051***	-0.0441***	-0.1149***		0.0123	0.0077***	0.0054***	-0.0022
	15sec	-0.0001*	0.0000***	-0.0000***	-0.0000***		0.0000	0.0000***	0.0000	0.0000
	30sec	-0.9628*	0.0039***	-0.0401***	-0.1101***		-0.0166	0.0089***	0.0014	-0.0071
	60sec	-0.9156*	0.0047***	-0.0373***	-0.1029***		-0.0638	0.0081***	-0.0013	-0.0142***
	300sec	-0.9360	0.0067***	-0.0305***	-0.0877***		-0.0433	0.0061***	-0.0081	-0.0294***
	600sec	-1.0064	0.0081***	-0.0275***	-0.0804***		0.0271	0.0046***	-0.0112	-0.0367***
	900sec	-0.8537	0.0085***	-0.0251***	-0.0770***		-0.1257	0.0043***	-0.0136*	-0.0401***
	1200sec	-0.8969	0.0082***	-0.0253***	-0.0766***		-0.0825	0.0046***	-0.0134	-0.0405***
	1500sec	-0.8836	0.0087***	-0.0244***	-0.0756***		-0.0958	0.0040***	-0.0142*	-0.0415***
	1800sec	-0.8715	0.0089***	-0.0246***	-0.0745***		-0.1079	0.0039**	-0.0141*	-0.0426***

*p<0.1; **p<0.05; ***p<0.01

Notes: This table describes the coefficients of the days to maturity ($\tilde{T}$) on a set of the volume weighted average market liquidity spreads for the 10-year full sample period, one year to maturity group ($\tilde{T}$ smaller than one year), three years to maturity group ($\tilde{T}$ smaller than three years), and five years to maturity group ($\tilde{T}$ smaller than five years). I utilize the following multivariate linear regression model to verify the maturity effects on the price discovery and market quality, $\tilde{\boldsymbol{Y}} = \boldsymbol{c} + \tilde{\boldsymbol{T}} + \boldsymbol{p^{-1}} + \boldsymbol{\mathcal{V}}$, where explanatory variables include the days to maturity ($\tilde{T}$), the inverse of transaction prices (p^{-1}), and the daily volatility ($\mathcal{V}$) is the highest daily transaction price subtracting the lowest daily transaction price. Dependent variables covers the volume weighted average bid-ask spreads ($\tilde{\mathscr{J}}^{Q}$), quoted half-spread ($\tilde{\mathscr{J}}^{Q^{1/2}}$), quoted depth ($\tilde{\mathscr{J}}^{D}$), effective half-spread ($\tilde{\mathscr{J}}^{E}$), realized spreads ($\tilde{\mathscr{J}}^{R}$), and adverse selection ($\tilde{\mathscr{J}}^{AS}$). Both realized spreads and adverse selection are measured at 26 varying time intervals, including tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec. It is worth noting that worths note that all liquidity spreads are reported in basis point (*bps*), except for the quoted depth. The days to maturity, the inverse of price and quoted depth are taken the logarithm value. All variables are the daily volume-weighted average results across five level order book data, with 32,712 observations in the 10-year full sample, 4,683 daily observations in the one year group, 15,009 daily observations in the three years group, and 25,481 daily observations in the five years group across 40 Eurodollar contracts. The asterisks $***$, $**$, and $*$ denote 1%, 5%, and 10% statistical significance of the coefficients of days to maturity. This table only reports the coefficient estimates of the days to maturity, other estimations and statistics are available in the online appendix.

large orders without impacting the prices.

It is worth noting that the one-year tenor coefficient estimations are all significantly positive for both effective spread and realized spread at different time intervals. This also indicates that the actual transaction prices are much closer to the mid-point price with more liquidity in the market when the contract approaches the maturity date. The tick-by-tick realized spread is the most affected by the days to maturity in year one group, with an increase of 0.2308 in basis points (see Table 5.12). Although the estimation of days to maturity is significant for the adverse selection at different timescales for the full sample, the regression results for three subsamples illustrate that the distance to maturity and adverse selection have a close relation. Except for the tick-by-tick adverse selection with -0.0121, others all show positively related to the days to maturity in the one-year to maturity subsample. This finding is consistent with the closer to maturity, the higher liquidity in the market. For the three-year to maturity and five-year to maturity subsamples, the coefficients are positive from tick level to 15 seconds time intervals, and turn to a negative relation if the time intervals are longer than one minute.

To evaluate the maturity effects on the adjustment of liquidity spreads, Table 5.13 also reports the estimation between spread autocorrelation at first-order, fifth-order, and tenth-order with the days to maturity. The days to maturity has negative impacts on the autocorrelation of both the bid-ask spread and quoted half spread, with the coefficients as -0.0515 in first-order, -0.0923 in fifth-order, -0.1080 in tenth-order. This means that the bid-ask spreads and quoted half spreads are increasingly autocorrelated when the contract moves to its maturity date. The price adjustment in facing the new information is likely to be less sensitive to lower price efficiency. Similarly, the negative coefficients for the auto-

Table 5.13: The First-order, Fifth-order and Tenth-order Spreads Autocorrelation Regression Results

		First-order				Fifth-order				Tenth-order			
		Full Sample	1-Year Tenor	3-Year Tenor	5-Year Tenor	Full Sample	1-Year Tenor	3-Year Tenor	5-Year Tenor	Full Sample	1-Year Tenor	3-Year Tenor	5-Year Tenor
$\Omega^{\bar{\mathscr{F}}Q}$		-0.0515***	-0.0026**	-0.0086***	-0.0227***	-0.0923***	-0.0086***	-0.0359***	-0.0525***	-0.1080***	-0.0127***	-0.0250***	-0.0686***
$\Omega^{\bar{\mathscr{F}}Q^{1/2}}$		-0.0515***	-0.0026**	-0.0086***	-0.0227***	-0.0923***	-0.0086***	-0.0359***	-0.0525***	-0.1080***	-0.0127***	-0.0250***	-0.0686***
$\Omega^{\bar{\mathscr{F}}D}$		-0.0392***	-0.0036	-0.0170***	-0.0140***	-0.0711***	-0.0162***	-0.0459***	-0.0328***	-0.0829***	-0.0218***	-0.0363***	-0.0427***
$\Omega^{\bar{\mathscr{F}}E}$		-0.0128***	-0.0387***	0.0072***	0.0108***	-0.0245***	-0.0103***	0.0216***	0.0065***	-0.0392***	0.0393***	0.0187***	-0.0077***
$\Omega^{\bar{\mathscr{F}}R}$	**Tick**	-0.0090***	-0.0355***	0.0113***	0.0153***	-0.0207***	-0.0064*	0.0251***	0.0115***	-0.0373***	0.0420***	0.0236***	-0.0045***
	1ms	-0.0129***	-0.0387***	0.0072***	0.0108***	-0.0245***	-0.0102***	0.0216***	0.0065***	-0.0393***	0.0393***	0.0186***	-0.0077***
	2ms	-0.0129***	-0.0386***	0.0072***	0.0108***	-0.0246***	-0.0102***	0.0215***	0.0065***	-0.0394***	0.0393***	0.0186***	-0.0078***
	5ms	-0.0131***	-0.0385***	0.0071***	0.0106***	-0.0248***	-0.0100***	0.0214***	0.0062***	-0.0396***	0.0394***	0.0185***	-0.0080***
	10ms	-0.0132***	-0.0382***	0.0072***	0.0105***	-0.0250***	-0.0097***	0.0213***	0.0060***	-0.0398***	0.0397***	0.0185***	-0.0084***
	15ms	-0.0133***	-0.0379***	0.0073***	0.0104***	-0.0252***	-0.0093***	0.0212***	0.0057***	-0.0400***	0.0398***	0.0185***	-0.0087***
	20ms	-0.0133***	-0.0376***	0.0073***	0.0103***	-0.0254***	-0.0090***	0.0211***	0.0054***	-0.0402***	0.0399***	0.0184***	-0.0089***
	25ms	-0.0135***	-0.0374***	0.0074***	0.0102***	-0.0254***	-0.0088***	0.0210***	0.0053***	-0.0403***	0.0400***	0.0184***	-0.0091***
	50ms	-0.0136***	-0.0364***	0.0076***	0.0100***	-0.0255***	-0.0079**	0.0205***	0.0049***	-0.0407***	0.0400***	0.0184***	-0.0098***
	75ms	-0.0134***	-0.0357***	0.0077***	0.0100***	-0.0256***	-0.0072**	0.0200***	0.0047***	-0.0411***	0.0400***	0.0183***	-0.0104***
	100ms	-0.0136***	-0.0351***	0.0078***	0.0098***	-0.0256***	-0.0066*	0.0196***	0.0045***	-0.0413***	0.0401***	0.0183***	-0.0108***
	150ms	-0.0135***	-0.0342***	0.0082***	0.0099***	-0.0255***	-0.0059*	0.0193***	0.0045***	-0.0414***	0.0403***	0.0185***	-0.0111***
	200ms	-0.0133***	-0.0338***	0.0086***	0.0102***	-0.0251***	-0.0055	0.0192***	0.0048***	-0.0413***	0.0402***	0.0189***	-0.0112***
	500ms	-0.0125***	-0.0318***	0.0096***	0.0113***	-0.0245***	-0.0039	0.0197***	0.0055***	-0.0406***	0.0405***	0.0200***	-0.0108***
	750ms	-0.0128***	-0.0304***	0.0102***	0.0113***	-0.0241***	-0.0025	0.0208***	0.0061***	-0.0401***	0.0413***	0.0212***	-0.0103***
	1sec	-0.0131***	-0.0292***	0.0109***	0.0112***	-0.0240***	-0.0015	0.0218***	0.0065***	-0.0400***	0.0416***	0.0223***	-0.0098***
	5sec	-0.0122***	-0.0221***	0.0160***	0.0136***	-0.0212***	0.0053	0.0289***	0.0106***	-0.0352***	0.0442***	0.0294***	-0.0041**
	15sec	-0.0107***	-0.0170***	0.0195***	0.0160***	-0.0191***	0.0085**	0.0306***	0.0130***	-0.0324***	0.0430***	0.0320***	-0.0018
	30sec	-0.0102***	-0.0142***	0.0210***	0.0171***	-0.0199***	0.0103***	0.0290***	0.0125***	-0.0334***	0.0428***	0.0319***	-0.0034**
	60sec	-0.0099***	-0.0125***	0.0220***	0.0181***	-0.0212***	0.0102***	0.0256***	0.0105***	-0.0359***	0.0396***	0.0304***	-0.0066***
	300sec	-0.0090***	-0.0069***	0.0252***	0.0213***	-0.0218***	0.0116***	0.0226***	0.0115***	-0.0374***	0.0361***	0.0308***	-0.0079***
	600sec	-0.0081***	-0.0063**	0.0260***	0.0225***	-0.0226***	0.0100***	0.0214***	0.0119***	-0.0373***	0.0328***	0.0303***	-0.0075***
	900sec	-0.0082***	-0.0063**	0.0267***	0.0227***	-0.0222***	0.0100***	0.0224***	0.0123***	-0.0375***	0.0314***	0.0314***	-0.0070***
	1200sec	-0.0071***	-0.0068***	0.0270***	0.0232***	-0.0225***	0.0084***	0.0215***	0.0121***	-0.0378***	0.0295***	0.0311***	-0.0077***
	1500sec	-0.0069***	-0.0075***	0.0272***	0.0233***	-0.0222***	0.0073**	0.0216***	0.0125***	-0.0374***	0.0280***	0.0313***	-0.0072***
	1800sec	-0.0074***	-0.0076***	0.0268***	0.0229***	-0.0221***	0.0072**	0.0209***	0.0122***	-0.0373***	0.0273***	0.0307***	-0.0074***
$\Omega^{\bar{\mathscr{F}}AS}$	**Tick**	-0.0187***	-0.0027	0.0251***	0.0011	-0.0237***	-0.0048	0.0004	-0.0079***	-0.0283***	-0.0013	0.0112***	-0.0161***
	1ms	0.1073***	0.0563***	0.1082***	0.1147***	0.0368***	0.0098**	0.0174***	0.0414***	0.0125***	0.0095***	0.0387***	0.0161***
	2ms	0.0835***	0.0325***	0.0783***	0.0923***	0.0185***	0.0087**	0.0103***	0.0250***	0.0020*	0.0109***	0.0264***	0.0069***
	5ms	0.0504***	0.0210***	0.0601***	0.0657***	0.0101***	0.0174***	0.0141***	0.0220***	-0.0017	0.0159***	0.0300***	0.0063***
	10ms	0.0581***	0.0292***	0.0750***	0.0795***	0.0105***	0.0224***	0.0194***	0.0272***	-0.0036***	0.0215***	0.0385***	0.0079***
	15ms	0.0599***	0.0372***	0.0822***	0.0846***	0.0136***	0.0263***	0.0243***	0.0333***	-0.0034**	0.0270***	0.0460***	0.0108***
	20ms	0.0612***	0.0397***	0.0854***	0.0877***	0.0165***	0.0296***	0.0281***	0.0390***	-0.0030**	0.0300***	0.0523***	0.0135***
	25ms	0.0623***	0.0387***	0.0874***	0.0901***	0.0189***	0.0336***	0.0321***	0.0436***	-0.0021	0.0328***	0.0582***	0.0160***
	50ms	0.0508***	0.0329***	0.0826***	0.0850***	0.0216***	0.0476***	0.0482***	0.0550***	-0.0029*	0.0526***	0.0759***	0.0225***
	75ms	0.0414***	0.0270***	0.0758***	0.0779***	0.0211***	0.0488***	0.0546***	0.0583***	-0.0047***	0.0591***	0.0811***	0.0244***
	100ms	0.0357***	0.0258***	0.0715***	0.0725***	0.0204***	0.0505***	0.0587***	0.0594***	-0.0057***	0.0643***	0.0838***	0.0253***
	150ms	0.0291***	0.0246***	0.0671***	0.0676***	0.0191***	0.0528***	0.0631***	0.0601***	-0.0067***	0.0714***	0.0851***	0.0265***
	200ms	0.0251***	0.0248***	0.0647***	0.0647***	0.0187***	0.0539***	0.0660***	0.0608***	-0.0069***	0.0746***	0.0861***	0.0278***
	500ms	0.0168***	0.0226***	0.0573***	0.0571***	0.0153***	0.0532***	0.0726***	0.0602***	-0.0065***	0.0787***	0.0847***	0.0319***
	750ms	0.0127***	0.0199***	0.0533***	0.0528***	0.0126***	0.0519***	0.0733***	0.0586***	-0.0071***	0.0792***	0.0827***	0.0327***
	1sec	0.0090***	0.0169***	0.0498***	0.0493***	0.0100***	0.0484***	0.0729***	0.0565***	-0.0079***	0.0762***	0.0799***	0.0327***
	5sec	-0.0032	0.0070**	0.0383***	0.0364***	-0.0035*	0.0368***	0.0602***	0.0417***	-0.0162***	0.0642***	0.0636***	0.0240***
	15sec	-0.0065***	0.0029	0.0345***	0.0309***	-0.0111***	0.0284***	0.0501***	0.0322***	-0.0225***	0.0550***	0.0533***	0.0155***
	30sec	-0.0085***	0.0011	0.0323***	0.0284***	-0.0162***	0.0254***	0.0425***	0.0262***	-0.0278***	0.0515***	0.0473***	0.0089***
	60sec	-0.0104***	-0.0003	0.0295***	0.0255***	-0.0207***	0.0217***	0.0341***	0.0197***	-0.0330***	0.0464***	0.0403***	0.0021
	300sec	-0.0133***	-0.0035	0.0255***	0.0212***	-0.0267***	0.0139***	0.0229***	0.0122***	-0.0398***	0.0345***	0.0313***	-0.0064***
	600sec	-0.0127***	-0.0019	0.0251***	0.0217***	-0.0285***	0.0116***	0.0198***	0.0108***	-0.0407***	0.0301***	0.0287***	-0.0078***
	900sec	-0.0125***	-0.0024	0.0253***	0.0215***	-0.0282***	0.0106***	0.0204***	0.0107***	-0.0410***	0.0283***	0.0290***	-0.0076***
	1200sec	-0.0119***	-0.0037	0.0253***	0.0215***	-0.0288***	0.0075**	0.0190***	0.0097***	-0.0418***	0.0255***	0.0281***	-0.0088***
	1500sec	-0.0117***	-0.0043*	0.0255***	0.0214***	-0.0284***	0.0065**	0.0186***	0.0101***	-0.0417***	0.0229***	0.0284***	-0.0088***
	1800sec	-0.0117***	-0.0044*	0.0253***	0.0213***	-0.0278***	0.0066**	0.0181***	0.0101***	-0.0413***	0.0227***	0.0279***	-0.0089***

* p<0.1; ** p¡0.05; *** p<0.01

Notes: This table describes the coefficients of the days to maturity ($\tilde{T}$) on a set of the volume weighted average spreads autocorrelations ($\Omega^{\bar{\mathscr{F}}}$) for the 10-year full sample period, one year to maturity group ($\tilde{T}$ smaller than one year), three years to maturity group ($\tilde{T}$ smaller than three years), and five years to maturity group ($\tilde{T}$ smaller than five years). I utilize the following multivariate linear regression model to verify the maturity effects on the price discovery and market quality, $\tilde{\mathbf{Y}} = \mathbf{c} + \tilde{\mathbf{T}} + \mathbf{p}^{-1} + \mathcal{V}$, where explanatory variables include the days to maturity ($\tilde{T}$), the inverse of transaction prices (p^{-1}), and the daily volatility ($\mathcal{V}$) is the highest daily transaction price subtracting the lowest daily transaction price. Dependent variables covers the volume weighted average first-order, fifth-order, and tenth-order autocorrelations of bid-ask spreads ($\Omega^{\bar{\mathscr{F}},Q}$), quoted half-spread ($\Omega^{\bar{\mathscr{F}},Q,1/2}$), quoted depth ($\Omega^{\bar{\mathscr{F}},D}$), effective half-spread ($\Omega^{\bar{\mathscr{F}},E}$), realized spreads ($\Omega^{\bar{\mathscr{F}},R}$), and adverse selection ($\Omega^{\bar{\mathscr{F}},AS}$). Both realized spreads autocorrelations and adverse selection autocorrelations are measured at 26 varying time intervals, including tick-by-tick, 1 ms, 2 ms, 5 ms, 10 ms, 15 ms, 20 ms, 25 ms, 50 ms, 75 ms, 100 ms, 150 ms, 200 ms, 500 ms, 750 ms, 1 sec, 5 sec, 15 sec, 30 sec, 60 sec, 300 sec, 600 sec, 900 sec, 1,200 sec, 1,500 sec, and 1,800 sec. It worth noting that the days to maturity and the inverse of price are taken the logarithm value. All variables are the daily volume-weighted average results across five level order book data, with 32,712 observations in the 10-year full sample, 4,683 daily observations in the one year group, 15,009 daily observations in the three years group, and 25,481 daily observations in the five years group across 40 Eurodollar contracts. The asterisks $***$, $**$, and $*$ denote 1%, 5%, and 10% statistical significance of the coefficients of days to maturity. This table only reports the coefficient estimates of the days to maturity, other estimations and statistics are available in the online appendix.

correlations of quoted depth, effective spreads and realized spreads of the whole sample indicates that, the closer to the maturity date, the higher the autocorrelation of these three spreads. The higher absolute value of the coefficients indicates a larger influence of the maturity effects within one year to maturity sample in the first-order autocorrelations. Except for the tick level, the coefficients on the first-order and fifth-order adverse selection autocorrelations are positive for the whole sample when the time intervals are less than one second. The adverse selection is less dependent on the previous one when it closes to maturity. However, the coefficient becomes negative as the time intervals are longer than one second. Above the one-second timescale, the first-order and fifth-order adverse selection autocorrelations increase when it moves to the settlement date.

5.5 Chapter Summary

I have provided a comprehensive empirical analysis of the high-frequency price discovery mechanism, the high-frequency trading performance and the maturity effects on market quality and price efficiency using a population study of the Eurodollar futures contracts. My dataset covers the entire market depth data and transactions from 2008 to 2014 at tick-by-tick level with millisecond timestamps, which contains every trade and every message update in the limited order-book, the whole dataset includes 2.63 billion quotes observations and 263.54 million trades observations. I believe that such a study on the high-frequency price discovery and market performance has never been attempted before.

To capture the efficiency of price incorporating information, I measure a broad range of first-order, fifth-order, and tenth-order mid-quoted prices and return autocorrelations with varying time intervals. Because I match the timestamps for

the mid-quoted return with the actual transaction, the results of return autocorrelations are much lower and more appropriate than the mid-price autocorrelations. The evidence implies that the mid-quoted return autocorrelations gradually increase from the shortest time interval to the longest time interval. Except for the tick-by-tick mid-quoted return autocorrelation, the other returns autocorrelations are all positive. The higher order of return autocorrelations is, the smaller return autocorrelations with the better price efficiency is. The first-order autocorrelations of volume weighted average returns are below 0.5 as the time interval shorter than 15 milliseconds. The shorter time interval is, the higher efficiency of price adjustment is.

Furthermore, this study also illustrates the use of my new volume weighted average trade classification algorithm that replaces the Lee-Ready algorithm to measure the trade direction within very short time intervals across the complete limit order book. As a contribution to this study, this new empirical technique has been employed to calculate a range of market liquidity spreads and spread autocorrelations. Moreover, the new trade direction indicator also contributes to the calculation of short-term pricing errors.

Following Hasbrouck [1991a,b, 1993] microstructure theory, the price can be decomposed into two parts: the long-term efficient price and the short-term pricing error. This chapter proxies the pricing error using the VAR model and adds both the bid-side and ask-side market concentration as variables. The impulse response of trade returns indicates that all the variables have the ability to influence the trade return at short time intervals. The largest response to trade return is from the shock to itself from the perspective of whole ten years contract life-cycle, followed by the response of trade return to the shock in the signed bid-side market concentration and the trade direction. Alternatively, the fastest adjustment of

trade return is also to itself, and finishes the whole adjustment process within 200 milliseconds. The evidence also suggests that there are critical turning points in the response of trade returns to the shocks among varying market indicators.

To provide a complementary understanding of the price discovery, I also assess the maturity effects on price efficiency and market quality in the Eurodollar future market. Overall, the days to maturity has a positive relation with the pricing errors, which indicates that the closer to maturity, the farther the distance between the actual transaction price and the efficient price. In other words, as the Eurodollar future approaches its maturity, the trade prices are less sensitive when incorporating available information in the market. This could be caused by the extremely actively traded Eurodollar contract accompanied by large amounts of information and inside trading activities. For the maturity effects with the mid-price return autocorrelations, the linear regression coefficients also demonstrate the days to maturity periodicity and significantly affect the price efficiency in the Eurodollar future market. More specifically, the distance to maturity has the largest impacts when the mid-quote return autocorrelation is measured at 1 ms time intervals.

Chapter 6

Conclusions

This thesis provides an extensive analysis of the microstructure of the Eurodollar future market and the impacts of high-frequency algorithmic trading activities on the quality of this market. As a hot topic, a large-scale studies have some mixed commentary on the impacts of HFT on financial markets characteristics. Recent work by Kyle and Obizhaeva [2013] and Foucault et al. [2015] present theoretical frameworks against HFT, and indicate that HFT simply through the faster access speed to benefit and brings negative effects to other traders. Hasbrouck [2014], Chakrabarty et al. [2014] and Kirilenko et al. [2014] demonstrate empirically that HFT increases the price volatility, trading costs, and damage the capacity of price discovery and the market liquidity. However, Hendershott et al. [2011], Riordan and Storkenmaier [2012] and Chaboud et al. [2014] comment on that HFT, as the liquidity provider, improve the information processing capacity of the financial markets and benefits the market efficiency and quality.

A three-month delivery Eurodollar future is a Libor-referenced contract with a unique IMM date. To understand the full picture of the microstructure of

Eurodollar future market, I analyze its limit order book with a population study of this future market. The data includes the entire limit order book from Jan 2008 to July 2014, and the inside quotes and transaction from Jan 1996 to July 2014. All data are gathered at tick-level with milliseconds timestamps. The thesis addresses the relation between HFT and the market characteristics, so the first step is to identify the market characteristics and to proxy the HFT, as I do not have the HFT identification or traders ID in my dataset. The next step is to develop a novel empirical methodology to capture the non-linear relation between the impact of the fraction of algorithmic trading and a large set of different market quality indicators. Last but not least, I measure the efficiency of price formation process on interest rate futures and the trading mechanism.

6.1 Concluding Remarks

Notes on Chapter 2

Chapter 2 analyzes the microstructure of the quotes volatility in the Eurodollar futures market, and the impact of high-frequency quoting on millisecond timescale volatility. The volatility is one of the core components of the market quality. Following the recently developed approach of Hasbrouck [2014], I adopt a Haar wavelet multi-resolution approach, a signal processing technique, to decompose an original price signal into a high-frequency domain and a low-frequency domain. Therefore, I can extract the volatility scaling (variance and covariance ratios) of the order flow over days at different timescales. The variance ratio, the price variation between high-frequency timescale and daily-level timescale, could be an indicator of high-frequency microstructure noise; the covariance ratio indicates

that the agreement between bid prices and ask prices. This chapter employs every inside-quote update (best bids and best asks) and tick-level transactions data between the beginning of 1996 and the end of 2013.

The findings in this chapter suggest, unlike the equity market, the variance ratio varies considerably with maturity. For the furthest tenors in the Eurodollar term structure, there is little trading activities, unlike the equity market, and the variance ratios of the inside quotes are well behaved and approach unity. However, both the trading activity and variance ratio increase substantially as the contracts approach maturity, especially two-year and one-year tenor contracts. This illustrates the high-frequency microstructure noise inside the bid-ask quoting prices which could cause the higher quotes variance at high frequency time-scales. The quoting volatility substantially declines as the contract approaches maturity (one or two weeks before its maturity date). Furthermore, this chapter illustrates the existence of high variance scaling and low covariance ratio at short timescales (millisecond level), which suggests the existence of high-frequency microstructure noises and the disagreement between ask and bid side at highest frequency in this market.

In short, as a Eurodollar future contract approaching its maturity, the quote prices become more volatile, and the high frequency trading increases the microstructure noises into the price volatility. My findings provide empirical evidence to support that the high-frequency trading activities do have effects on the maturity effects of this type future market, especially for the two to one year tenor Eurodollar contracts. It also points out that HFT shifts trading behavior since the Eurodollar contract passes its maturity. The identification of this critical time, maybe help to those seeking to avoid direct trading against HFTs.

Notes on Chapter 3

Chapter 3 provides a standard asymmetric information based theoretical model to predict the relation on the term structure of Eurodollar future contracts. The main theoretical prediction is a non-linear relation between the saturation of AT versus the impacts on the quality of the market. This chapter also covers a broad battery of liquidity provision measurements and provides the measurements for the high frequency algorithmic trading based on the quotes messages updating speed. To proxy the fraction of high-frequency algorithmic trading, it requires to have every quoting update in the market, hence, the data utilized in this chapter move from the inside quotes to the entire limit order book from Jan 2008 to July 2014.

I demonstrate the different liquidity measurements are consistent over the term structure of Eurodollar contracts. The contract liquidity of the initial three years is relatively stable; however, as the contract moves close to maturity, the spreads get narrower and more volatile among the five levels (the quoted depth become deeper). Particularly within two-year tenor, spreads become much narrower indicating greater market liquidity. Moreover, the results demonstrate level two quotes provide the largest liquidity into the market with the highest depth.

Then, I generate the approximation of the proportion of high frequency algorithmic trading for both the bid and ask sides with a set of thresholds from 25 milliseconds to 200 milliseconds. This chapter explores that the majority of quotes under 200ms. For the proxy thresholded at 25 ms, the most frequently proportion of algorithmic activities in this market is around 65%, which is consistent with the official report from CME Group [2010]. It is worth mentioning that the second level is the most rapidly and actively messaged among the entire

order book.

Notes on Chapter 4

This chapter captures a non-linear relation between the fraction of high-speed algorithmic trading and market quality via a novel partially linear semi-parametric regression estimator. My empirical results show the similar pattern and are consistent with the theoretical prediction. Overall, this chapter suggests that the high-frequency algorithmic trading does have impacts on the market quality, and thresholds (or saturation points) exist in terms of the impact of HFTs on the execution risks. Specifically, when there is a small number of algorithmic traders in the market, AT do have impacts on the execution risk and mainly in the market liquidity. When the proportion of AT increases, the execution costs are primarily from the price volatility. Once the AT proxy reaches a certain saturation point, the marginal effect on execution risks seems to disappear. This finding may partially explain the puzzle why prior studies have found the contradictory evidence regarding the impact of HFTs on market characteristics.

Notes on Chapter 5

Chapter 5 evaluates the capacity of the high-frequency price formation and discovery mechanism using the tick level limited order book data from 2008 to 2014. This chapter tells you how efficient of prices incorporating information in the market and how fast the Eurodollar future prices can adjust to the efficient prices. In this chapter, I measure a broad range of quotes and trades autocorrelations at first-order, fifth-order, and tenth-order at 26 timescales to capture the efficiency

of the Eurodollar future prices. The findings demonstrate that the mid-quoted return autocorrelation are positive and gradually increase from the high-frequency time-interval to the low-frequency time-interval. If I set 0.5 as the benchmark, the mid-quote prices can adjust to the information within 15 milliseconds. The shorter time interval is, the higher efficiency of mid-price adjustment is.

In addition, this chapter provides a new volume weighted average algorithm to classify the trades directions and employs its indicators in the following analysis. Because the traditional Lee-Ready algorithm lacks the accuracy dealing with tick data at millisecond timestamps. So I develop this novel approach specific for the microstructure studies utilizing order book data. This is another contribution of this thesis.

I then utilize a vector autoregression to capture the pricing error proxy, as the deviation of transaction prices from the efficient prices to illustrate the term structure of these costs. The impulse response of trade returns also reports that all the factors (the trade return, signed trade direction, signed trade volume, signed square root of trade volume, and signed asks and bids market concentration ratio) have the ability to affect the trade return at very short time intervals. The fastest adjustment of trade returns is finished within 200 milliseconds from the shock to itself. It is worth noting that the three-year tenor contract has the fastest trade price formation process than one-year tenor and fifth-year tenor contracts.

This chapter also conducts a multivariate linear regression to capture the maturity effects on the Eurodollar future price formation process. The results illustrate that the distance to maturity can influence the price efficiency. In other words, as the Eurodollar future approaches its maturity, the trade prices are less sensitive when incorporating available information in the market.

6.2 Future Research

There are still some limitations in this thesis, for instance the measurements of high frequency trading. Although 25 ms interval is very fast, faster than human reaction, it is still possible that some accidental orders submit by humans located within this 25 milliseconds interval. So in the future work, I will consider improving the measurements of AT proxy. Meanwhile, I will try to simulate the Eurodollar market to see whether it is possible that a large amount of traders could it create the false appearance of HFT.

Secondly, because of the unique IMM date of this future market, one of the future work may look at the impact of trading activity on the term structure of Eurodollar futures market. Especially after the Dodd-Frank Act of 2010, trading in OTC-IRS interest rate swaps has become more heavily regulated and there has been a significant transfer of activity from the opaque OTC-IRS to using Eurodollar strips to construct synthetic swaps. With the exploded trading activities on the Eurodollar future market, it could be interesting to consider the changes in trading activity around IMM dates with a distorting effect on the benchmark.

Thirdly, because I have the full limited order book, in future, I will try to identify a set of malicious trading in the Eurodollar futures and assess its impacts on market quality. For instance, the iceberg order or hidden order is the order only report its prices but hidden parts of the quote volume.

One more thing is to extend the data period from middle of 2014 to the latest available date. It is worth noting that the European central bank has set the negative deposit rate in September 2014, which could have a significant impact on the Eurodollar deposits and its future market. My current data span, from 1996 to July 2014, does not cover the period after negative rate. It could be

interesting to dig a little deeper on the microstructure changes after the middle of 2014.

Last but not least, two techniques have developed in this thesis, one is the trade classification algorithmic to identify a buyer-initiated trade or seller-initiated trade, the other is a non-parametric estimator to determine the influence of algorithmic trading on the execution risks. So in future work, I will also focus on the development of these two techniques and try to apply them into the market microstructure research area.

Bibliography

Admati, A. R. (1985). A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica 53*(3), 629–657. 28, 83, 86, 92

Admati, A. R. and P. Pfleiderer (1988). A theory of intraday patterns: Volume and price variability. *The Review of Financial Studies 1*(1), 3–40. 82, 86, 88, 89, 92

Admati, A. R. and P. Pfleiderer (1989). Divide and conquer: A theory of intraday and day-of-the-week mean effects. *The Review of Financial Studies 2*(2), 189–223. 82, 86, 92

Anderson, R. M., K. S. Eom, S. B. Hahn, and J.-H. Park (2013). Autocorrelation and partial price adjustment. *Journal of Empirical Finance 24*, 78–93. 177, 180, 181, 187, 216, 217

Anderson, R. W. and J.-P. Danthine (1983). The time pattern of hedging and the volatility of futures prices. *The Review of Economic Studies 50*(2), 249–266. 40

Angel, J. J., L. E. Harris, and C. S. Spatt (2011). Equity trading in the 21st century. *The Quarterly Journal of Finance 1*(1), 1–53. 99, 100

Baillie, R. T., G. G. Booth, Y. Tse, and T. Zabotina (2002). Price discovery and common factor models. *Journal of Financial Markets 5*(3), 309–321. 180

Barclay, M. J. and T. Hendershott (2003). Price discovery and trading after hours. *The Review of Financial Studies 16*(4), 1041–1073. 183

Bershova, N. and D. Rakhlin (2013). High-frequency trading and long-term investors: A view from the buy-side. *Journal of Investment Strategies 2*(2), 25–69. 99

Bessembinder, H., J. F. Coughenour, P. J. Seguin, and M. M. Smoller (1996). Is there a term structure of futures volatilities? reevaluating the Samuelson hypothesis. *The Journal of Derivatives 4*(2), 45–58. 40, 60, 204, 240

Bessembinder, H. and K. Venkataraman (2010). Bid-ask spreads: Measuring trade execution costs in financial markets. In R. Cont (Ed.), *Encyclopedia of Quantitative Finance*. John Wiley & Sons, New York, NY. 179

Beveridge, S. and C. R. Nelson (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics 7*(2), 151–174. 200

Biais, B. and P.-O. Weill (2009). Liquidity shocks and order book dynamics. Working Paper. 100

Bianco, S. and R. Renò (2006). Dynamics of intraday serial correlation in the italian futures market. *Journal of Futures Markets 26*(1), 61–84. 217

Boehmer, E. and E. K. Kelley (2009). Institutional investors and the informational efficiency of prices. *The Review of Financial Studies 22*(9), 3563–3594. 183

Boehmer, E. and J. J. Wu (2013). Short selling and the price discovery process. *The Review of Financial Studies 26*(2), 287–322. 183, 203

Bollen, B. and B. Inder (2002). Estimating daily volatility in financial markets utilizing intraday data. *Journal of Empirical Finance 9*(5), 551–562. 40

Brigo, D. and F. Mercurio (2006). *Interest rate models-theory and practice: with smile, inflation and credit* (2nd ed.). Springer. 27, 30, 32

British Bankers' Association (2008). Understanding the construction and operation of BBA LIBOR-strengthening for the future. Consultative paper June 10. 29

Brogaard, J. (2010). High frequency trading and its impact on market quality. Working Paper, Northwestern University Kellogg School of Management. 100

Brogaard, J. (2012). *Essays on High Frequency Trading*. Ph. D. thesis, Northwestern University. 39

Brogaard, J., T. Hendershott, and R. Riordan (2014). High-frequency trading and price discovery. *The Review of Financial Studies 27*(8), 2267–2306. 39, 180, 184, 185

Budish, E. B., P. Cramton, and J. J. Shim (2013). The high-frequency trading arms race: Frequent batch auctions as a market design response. Fama-Miller Working Paper. 148

Carrion, A. (2013). Very fast money: High-frequency trading on the nasdaq. *Journal of Financial Markets 16*(4), 680–711. 99, 179

Castura, J., R. Litzenberger, R. Gorelick, and Y. Dwivedi (2010). Market efficiency and microstructure evolution in US equity markets: A high-frequency perspective. Working Paper, RGM Advisors, LLC. 3, 99, 100, 180, 184

Chaboud, A. P., B. Chiquoine, E. Hjalmarsson, and C. Vega (2014). Rise of the machines: Algorithmic trading in the foreign exchange market. *The Journal of Finance 69*(5), 2045–2084. 3, 4, 97, 184, 185, 252

Chakrabarty, B., P. K. Jain, A. Shkilko, and K. Sokolov (2014). Speed of market access and market quality: Evidence from the sec naked access ban. Working Paper, Western Finance Association 2014. 3, 177, 183, 185, 186, 252

Chan, K. (1992). A further analysis of the lead-lag relationship between the cash market and stock index futures market. *The Review of Financial Studies 5*(1), 123–152. 4, 96, 184

Chen, J. and X. Zheng (2013). Price volatility and contract maturity: Evidence from an online futures market for sports tickets. *Eastern Economic Journal 40*(1), 56–70. 204

Chen, Y.-J., J.-C. Duan, and M.-W. Hung (1999). Volatility and maturity effects in the nikkei index futures. *Journal of Futures Markets 19*(8), 895–909. 40

CME Group (2010). Algorithmic trading and market dynamics. Technical report, CME Group. 10, 79, 97, 105, 113, 114, 119, 120, 122, 176, 255

Commodity Futures Trading Commission (2012, June). In the matter of barclays plc, barclays bank plc, and barclays capital inc., order instituting proceedings pursuant to sections 6(c) and 6(d) of the commodity exchange act, as amended, making findings and imposing remedial sanctions. CFTC Docket No 12–25. 4, 19, 20

Conlon, T. and J. Cotter (2012). An empirical analysis of dynamic multiscale hedging using wavelet decomposition. *Journal of Futures Markets 32*(3), 272–299. 42

Conlon, T., M. Crane, and H. J. Ruskin (2008). Wavelet multiscale analysis for hedge funds: Scaling and strategies. *Physica A: Statistical Mechanics and its Applications 387*(21), 5197–5204. 42

Daal, E., J. Farhat, and P. P. Wei (2006). Does futures exhibit maturity effect? new evidence from an extensive set of US and foreign futures contracts. *Review of Financial Economics 15*(2), 113–128. 40

Darolles, S., Y. Fan, J.-P. Florens, and E. Renault (2011). Nonparametric instrumental regression. *Econometrica 79*(5), 1541–1565. 133, 135

Domowitz, I. and H. Yegerman (Winter 2006). The cost of algorithmic trading a first look at comparative performance. *The Journal of Trading 1*(1), 33–42. 97

Dufour, A. and R. F. Engle (2000). Time and the price impact of a trade. *The Journal of Finance 55*(6), 2467–2498. 183

Engle, R. and A. Patton (2004). Impacts of trades in an error-correction model of quote prices. *Journal of Financial Markets 7*(1), 1–25. 183

Engle, R. F., R. Ferstenberg, and J. Russell (2012). Measuring and modeling execution cost and risk. *Journal of Portfolio Management 38*(2), 14–28. 4, 96

Eom, K. S., S. B. Hahn, and S. Joo (2004). Partial price adjustment and autocorrelation in foreign exchange markets. Working Paper, University of California at Berkeley. 180, 187

Florens, J.-P., J. Johannes, and S. Van Bellegem (2012). Instrumental regression in partially linear models. *The Econometrics Journal 15*(2), 304–324. 131, 132, 133, 135, 138

Foley, S. and T. J. Putniņš (2014). Should we be afraid of the dark? dark trading and market quality. Working Paper. 180, 181, 182, 187, 216, 217

Foucault, T., J. Hombert, and I. Roşu (2015). News trading and speed. *The Journal of Finance 71*(1), 335–382. 3, 252

Frino, A. and M. D. McKenzie (2002). The impact of screen trading on the link between stock index and stock index futures prices: Evidence from uk markets. European Financial Management Association (EFMA) 2002 London Meetings. 4, 96, 184

Gabor, D. (1946). Theory of communication. *Journal of the Institution of Electrical Engineers 93*, 429–457. 43

Galagedera, D. T. U. and E. A. Maharaj (2008). Wavelet timescales and conditional relationship between higher-order systematic co-moments and portfolio returns. *Quantitative Finance 8*(2), 201–215. 42

Galloway, T. M. and R. W. Kolb (1996). Futures prices and the maturity effect. *Journal of Futures Markets 16*(7), 809–828. 40

Gencay, R., F. Selcuk, and B. Whitcher (2002). An introduction to wavelets and other filtering methods in finance and economics. *Waves in Random Media 12*(3), 399–399. 42

Glosten, L. R. and P. R. Milgrom (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*(1), 71–100. 28, 83

Grammatikos, T. and A. Saunders (1986). Futures price variability: a test of maturity and volume effects. *The Journal of Business 59*(2), 319–330. 40

Grünbichler, A., F. A. Longstaff, and E. S. Schwartz (1994). Electronic screen trading and the transmission of information: An empirical examination. *Journal of Financial Intermediation 3*(2), 166–187. 4, 184

Gurrola, P. and R. Herrerías (2011). Maturity effects in the mexican interest rate futures market. *Journal of Futures Markets 31*(4), 371–393. 40

Haar, A. (1910). Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen 69*(3), 331–371. 43

Hagströmer, B. and L. Nordén (2013). The diversity of high-frequency traders. *Journal of Financial Markets 16*(4), 741–770. 39

Hall, P., J. L. Horowitz, et al. (2005). Nonparametric methods for inference in the presence of instrumental variables. *The Annals of Statistics 33*(6), 2904–2929. 134, 135

Hamilton, J. D. (1994). *Time series analysis*, Volume 2. Princeton university press, Princeton. 201

Han, L.-M., J. L. Kling, and C. W. Sell (1999). Foreign exchange futures volatility: Day-of-the-week, intraday, and maturity patterns in the presence of macroeconomic announcements. *Journal of Futures Markets 19*(6), 665–693. 40

Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics 24*(2), 127–161. 61

Harris, L. (1989). S&P 500 cash stock price volatilities. *The Journal of Finance 44*(5), 1155–1175. 4, 96, 184

Harris, L. (1990). Liquidity, trading rules and electronic trading systems. Technical report, Southern California-School of Business Administration. 98

Hasbrouck, J. (1991a). Measuring the information content of stock trades. *The Journal of Finance 46*(1), 179–207. 182, 183, 186, 198, 250

Hasbrouck, J. (1991b). The summary informativeness of stock trades: An econometric analysis. *The Review of Financial Studies 4*(3), 571–595. 182, 183, 186, 198, 250

Hasbrouck, J. (1993). Assessing the quality of a security market: A new approach to transaction-cost measurement. *The Review of Financial Studies 6*(1), 191–212. 181, 182, 183, 186, 198, 199, 200, 203, 242, 250

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The Journal of Finance 50*(4), 1175–1199. 180

Hasbrouck, J. (2014). High frequency quoting: Short-term volatility in bids and offers. Working Paper, Stern School of Business New York University. 2, 10, 19, 21, 25, 41, 42, 45, 46, 54, 60, 61, 62, 67, 70, 73, 127, 252, 253

Hendershott, T. and C. M. Jones (2005). Island goes dark: Transparency, fragmentation, and regulation. *The Review of Financial Studies 18*(3), 743–793. 177, 180, 181, 186, 187, 216, 217, 242

Hendershott, T., C. M. Jones, and A. J. Menkveld (2011). Does algorithmic trading improve liquidity? *The Journal of Finance 66*(1), 1–33. 3, 38, 62, 80, 97, 99, 100, 101, 102, 127, 179, 190, 197, 245, 252

Hendershott, T. and P. C. Moulton (2011). Automation, speed, and stock market quality: The nyse's hybrid. *Journal of Financial Markets 14*(4), 568–604. 3, 38

Hendershott, T. and R. Riordan (2011). Algorithmic trading and information. Working paper, University of California, Berkeley . 4, 184, 185

Hendershott, T. and R. Riordan (2013). Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis 48*(4), 1001–1024. 100, 101

Hong, H. (2000). A model of returns and trading in futures markets. *The Journal of Finance 55*(2), 959–988. 40

Hou, K. and T. J. Moskowitz (2005). Market frictions, price delay, and the cross-section of expected returns. *The Review of Financial Studies 18*(3), 981–1020. 181

Huang, R. D. and H. R. Stoll (1994). Market microstructure and stock return predictions. *The Review of Financial Studies 7*(1), 179–213. 4, 96, 184

Jarnecic, E. and M. Snape (2014). The provision of liquidity by high-frequency participants. *Financial Review 49*(2), 371–394. 100

Jovanovic, B. and A. J. Menkveld (2010). Middlemen in limit order markets. 2010 Meeting Papers 955, Society for Economic Dynamics. 99, 179

Kalev, P. S. and H. N. Duong (2008). A test of the samuelson hypothesis using realized range. *Journal of Futures Markets 28*(7), 680–696. 40, 41, 60, 204, 240

Karatzas, I. and C. Kardaras (2007). The numéraire portfolio in semimartingale financial models. *Finance and Stochastics 11*(4), 447–493. 32

Khoury, N. and P. Yourougou (1993). Determinants of agricultural futures price volatilities: Evidence from winnipeg commodity exchange. *Journal of Futures Markets 13*(4), 345–356. 40

Kirilenko, A., A. S. Kyle, M. Samadi, and T. Tuzun (2014). The flash crash: The impact of high frequency trading on an electronic market. Working paper. 3, 252

Kosinski, R. J. (2008). A literature review on reaction time. Technical report, Clemson University. 106

Krishnamurti, C., J. M. Sequeira, and F. Fangjian (2003). Stock exchange governance and market quality. *Journal of Banking & Finance 27*(9), 1859 – 1878. The future of stock exchanges in a globalizing world. 183

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica 53*(6), 1315–1335. 28, 83, 98

Kyle, A. S. and A. A. Obizhaeva (2013). Market microstructure invariance: Theory and empirical tests. Working paper, University of Maryland. 3, 252

Lee, C. M. C. and M. J. Ready (1991). Inferring trade direction from intraday data. *The Journal of Finance 46*(2), 733–746. 12, 102, 190, 192

Lehmann, B. N. (2002). Some desiderata for the measurement of price discovery across markets. *Journal of Financial Markets 5*(3), 259–276. 180

Levi, M. D. (2007). *International Finance: Contemporary Issues*. Routledge. 27

Lien, D. and K. Shrestha (2007). An empirical analysis of the relationship between hedge ratio and hedging horizon using wavelet analysis. *Journal of Futures Markets 27*(2), 127–150. 42

Lo, A. W. and A. C. MacKinlay (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *The Review of Financial Studies 1*(1), 41–66. 180, 184, 217

Madhavan, A. (2000). Market microstructure: A survey. *Journal of Financial Markets 3*(3), 205–258. 98

Malinova, K., A. Park, and R. Riordan (2013). Do retail traders benefit from improvements in liquidity? Working Paper, University of Toronto. 99

Menkveld, A. J. (2013). High frequency trading and the new market makers. *Journal of Financial Markets 16*(4), 712 – 740. 99

Menkveld, A. J. and M. A. Zoican (2014). Need for speed? exchange latency and market quality. Working paper. 4, 99, 179

Meyer, Y. (1993). *Wavelets-algorithms and applications*. SIAM, Philadelphia. 44, 46

Milonas, N. T. (1986). Price variability and the maturity effect in futures markets. *Journal of Futures Markets 6*(3), 443–460. 40

Musiela, M. and M. Rutkowski (2005). *Martingale methods in financial modelling* (2nd ed.), Volume 36 of *Stochastic Modelling and Applied Probability*. Springer. 27

Newey, W. K. and J. L. Powell (2003). Instrumental variable estimation of non-parametric models. *Econometrica 71*(5), 1565–1578. 131, 134

O'Hara, M. (1995). *Market microstructure theory*. Blackwell Cambridge, MA. 98

O'Hara, M. (2003). Presidential address: Liquidity and price discovery. *The Journal of Finance 58*(4), 1335–1354. 175

O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics 116*(2), 257 – 270. 190

Percival, D. B. and H. O. Mofjeld (1997). Analysis of subtidal coastal sea level fluctuations using wavelets. *Journal of the American Statistical Association 92*(439), 868–880. 42

Percival, D. B. and A. T. Walden (2000). *Wavelet Methods for Time Series Analysis*. Cambridge: Cambridge University Press. 42

Pilbeam, K. (2005). *International finance* (3rd ed.). Palgrave Macmillan. 27

Riordan, R. and A. Storkenmaier (2012). Latency, liquidity and price discovery. *Journal of Financial Markets 15*(4), 416 – 437. 3, 38, 175, 252

Robinson, P. M. (1988). Root-n-consistent semiparametric regression. *Econometrica 56*(4), 931–954. 131, 133

Rua, A. and L. C. Nunes (2009). International comovement of stock market returns: A wavelet analysis. *Journal of Empirical Finance 16*(4), 632–639. 42

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial management review 6*(2), 41–49. 40, 63, 64, 240

Schreiber, P. S. and R. A. Schwartz (1986). Price discovery in securities markets. *The Journal of Portfolio Management 12*(4), 43–48. 180

Stoll, H. R. and R. E. Whaley (1990). The dynamics of stock index and stock index futures returns. *Journal of Financial and Quantitative Analysis 25*(4), 441–468. 4, 96, 184

Watanabe, M. (2008). A model of stochastic liquidity. Technical Report 03-18, Yale ICF Working Paper. 85, 86, 88, 89, 92, 93

Yu, Y. and D. Ruppert (2002). Penalized spline estimation for partially linear single-index models. *Journal of the American Statistical Association 97*(460), 1042–1054. 137

Zhang, S. and R. Riordan (2011). Technology and market quality: the case of high frequency trading. In *ECIS 2011*. 99